

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Israel, David
Wolfman, Neil M
- (ii) TITLE OF INVENTION: RECOMBINANT BONE MORPHOGENETIC PROTEIN
HETERODIMERS, COMPOSITIONS AND METHODS OF USE
- (iii) NUMBER OF SEQUENCES: 35
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Legal Affairs, Genetics Institute, Inc.
 - (B) STREET: 87 CambridgePark Drive
 - (C) CITY: Cambridge
 - (D) STATE: MA
 - (E) COUNTRY: USA
 - (F) ZIP: 02140-2387
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 07/864,692
 - (B) FILING DATE: 07-APR-1992
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Kapinos, Ellen J.
 - (B) REGISTRATION NUMBER: 32,245
 - (C) REFERENCE/DOCKET NUMBER: GI-5192A
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 617 876-1170
 - (B) TELEFAX: 617 876-5851

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1607 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: DNA
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 356..1543
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GTCTGACTCTA GAGTGTGTGT CAGCTTTGG CTGGGGACTT CTTGAACTTCAGGGAGAAT	60
AACTTGCGCA CCCCACCTTG CGCCGGTGCC TTTGCCCCAG CGGAGCCTGC TTCGCCATCT	120
CCGAGCCCCA CCGCCCCTCC ACTCCTCGGC CTTGCCCCGAC ACTGAGACGC TGTTCCCAGC	180
GTGAAAAGAG AGACTGCGCG GCCGGCACCC GGGAGAAGGA GGAGGCAAAG AAAAGGAACG	240
GACATTCGGT CCTTGCGCCA GGTCCTTTGA CCAGAGTTTT TCCATGTGGA CGCTCTTTCA	300
ATGGACGTGT CCCC GCGTGC TTCTTAGACG GACTGCGGTC TCCTAAAGGT CGACC ATG Met 1	358
GTG GCC GGG ACC CGC TGT CTT CTA GCG TTG CTG CTT CCC CAG GTC CTC Val Ala Gly Thr Arg Cys Leu Leu Ala Leu Leu Leu Pro Gln Val Leu 5 10 15	406
CTG GGC GGC GCG GCT GGC CTC GTT CCG GAG CTG GGC CGC AGG AAG TTC Leu Gly Gly Ala Ala Gly Leu Val Pro Glu Leu Gly Arg Arg Lys Phe 20 25 30	454
GCG GCG GCG TCG TCG GGC CGC CCC TCA TCC CAG CCC TCT GAC GAG GTC Ala Ala Ala Ser Ser Gly Arg Pro Ser Ser Gln Pro Ser Asp Glu Val 35 40 45	502
CTG AGC GAG TTC GAG TTG CGG CTG CTC AGC ATG TTC GGC CTG AAA CAG Leu Ser Glu Phe Glu Leu Arg Leu Leu Ser Met Phe Gly Leu Lys Gln 50 55 60 65	550
AGA CCC ACC CCC AGC AGG GAC GCC GTG GTG CCC CCC TAC ATG CTA GAC Arg Pro Thr Pro Ser Arg Asp Ala Val Val Pro Pro Tyr Met Leu Asp 70 75 80	598
CTG TAT CGC AGG CAC TCA GGT CAG CCG GGC TCA CCC GCC CCA GAC CAC Leu Tyr Arg Arg His Ser Gly Gln Pro Gly Ser Pro Ala Pro Asp His 85 90 95	646
CGG TTG GAG AGG GCA GCC AGC CGA GCC AAC ACT GTG CGC AGC TTC CAC Arg Leu Glu Arg Ala Ala Ser Arg Ala Asn Thr Val Arg Ser Phe His 100 105 110	694
CAT GAA GAA TCT TTG GAA GAA CTA CCA GAA ACG AGT GGG AAA ACA ACC His Glu Glu Ser Leu Glu Glu Leu Pro Glu Thr Ser Gly Lys Thr Thr 115 120 125	742
CGG AGA TTC TTC TTT AAT TTA AGT TCT ATC CCC ACG GAG GAG TTT ATC Arg Arg Phe Phe Phe Asn Leu Ser Ser Ile Pro Thr Glu Glu Phe Ile 130 135 140 145	790
ACC TCA GCA GAG CTT CAG GTT TTC CGA GAA CAG ATG CAA GAT GCT TTA Thr Ser Ala Glu Leu Gln Val Phe Arg Glu Gln Met Gln Asp Ala Leu 150 155 160	838
GGA AAC AAT AGC AGT TTC CAT CAC CGA ATT AAT ATT TAT GAA ATC ATA Gly Asn Asn Ser Ser Phe His His Arg Ile Asn Ile Tyr Glu Ile Ile 165 170 175	886
AAA CCT GCA ACA GCC AAC TCG AAA TTC CCC GTG ACC AGA CTT TTG GAC Lys Pro Ala Thr Ala Asn Ser Lys Phe Pro Val Thr Arg Leu Leu Asp	934

180

185

190

ACC	AGG	TTG	GTG	AAT	CAG	AAT	GCA	AGC	AGG	TGG	GAA	ACT	TTT	GAT	GTC	982
Thr	Arg	Leu	Val	Asn	Gln	Asn	Ala	Ser	Arg	Trp	Glu	Thr	Phe	Asp	Val	
	195					200					205					
ACC	CCC	GCT	GTG	ATG	CGG	TGG	ACT	GCA	CAG	GGA	CAC	GCC	AAC	CAT	GGA	1030
Thr	Pro	Ala	Val	Met	Arg	Trp	Thr	Ala	Gln	Gly	His	Ala	Asn	His	Gly	
	210				215					220					225	
TTC	GTG	GTG	GAA	GTG	GCC	CAC	TTG	GAG	GAG	AAA	CAA	GGT	GTC	TCC	AAG	1078
Phe	Val	Val	Glu	Val	Ala	His	Leu	Glu	Glu	Lys	Gln	Gly	Val	Ser	Lys	
				230					235					240		
AGA	CAT	GTT	AGG	ATA	AGC	AGG	TCT	TTG	CAC	CAA	GAT	GAA	CAC	AGC	TGG	1126
Arg	His	Val	Arg	Ile	Ser	Arg	Ser	Leu	His	Gln	Asp	Glu	His	Ser	Trp	
			245					250					255			
TCA	CAG	ATA	AGG	CCA	TTG	CTA	GTA	ACT	TTT	GGC	CAT	GAT	GGA	AAA	GGG	1174
Ser	Gln	Ile	Arg	Pro	Leu	Leu	Val	Thr	Phe	Gly	His	Asp	Gly	Lys	Gly	
		260					265					270				
CAT	CCT	CTC	CAC	AAA	AGA	GAA	AAA	CGT	CAA	GCC	AAA	CAC	AAA	CAG	CGG	1222
His	Pro	Leu	His	Lys	Arg	Glu	Lys	Arg	Gln	Ala	Lys	His	Lys	Gln	Arg	
	275					280					285					
AAA	CGC	CTT	AAG	TCC	AGC	TGT	AAG	AGA	CAC	CCT	TTG	TAC	GTG	GAC	TTC	1270
Lys	Arg	Leu	Lys	Ser	Ser	Cys	Lys	Arg	His	Pro	Leu	Tyr	Val	Asp	Phe	
	290				295					300					305	
AGT	GAC	GTG	GGG	TGG	AAT	GAC	TGG	ATT	GTG	GCT	CCC	CCG	GGG	TAT	CAC	1318
Ser	Asp	Val	Gly	Trp	Asn	Asp	Trp	Ile	Val	Ala	Pro	Pro	Gly	Tyr	His	
				310				315						320		
GCC	TTT	TAC	TGC	CAC	GGA	GAA	TGC	CCT	TTT	CCT	CTG	GCT	GAT	CAT	CTG	1366
Ala	Phe	Tyr	Cys	His	Gly	Glu	Cys	Pro	Phe	Pro	Leu	Ala	Asp	His	Leu	
			325					330					335			
AAC	TCC	ACT	AAT	CAT	GCC	ATT	GTT	CAG	ACG	TTG	GTC	AAC	TCT	GTT	AAC	1414
Asn	Ser	Thr	Asn	His	Ala	Ile	Val	Gln	Thr	Leu	Val	Asn	Ser	Val	Asn	
		340					345					350				
TCT	AAG	ATT	CCT	AAG	GCA	TGC	TGT	GTC	CCG	ACA	GAA	CTC	AGT	GCT	ATC	1462
Ser	Lys	Ile	Pro	Lys	Ala	Cys	Cys	Val	Pro	Thr	Glu	Leu	Ser	Ala	Ile	
	355					360					365					
TCG	ATG	CTG	TAC	CTT	GAC	GAG	AAT	GAA	AAG	GTT	GTA	TTA	AAG	AAC	TAT	1510
Ser	Met	Leu	Tyr	Leu	Asp	Glu	Asn	Glu	Lys	Val	Val	Leu	Lys	Asn	Tyr	
	370				375					380					385	
CAG	GAC	ATG	GTT	GTG	GAG	GGT	TGT	GGG	TGT	CGC	TAGTACAGCA AAATTAAATA				1563	
Gln	Asp	Met	Val	Val	Glu	Gly	Cys	Gly	Cys	Arg						
				390				395								
CATAAATATA TATATATATA TATATTTTAG AAAAAAGAAA AAAA																1607

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Val	Ala	Gly	Thr	Arg	Cys	Leu	Leu	Ala	Leu	Leu	Leu	Pro	Gln	Val	
1				5					10					15		
Leu	Leu	Gly	Gly	Ala	Ala	Gly	Leu	Val	Pro	Glu	Leu	Gly	Arg	Arg	Lys	
			20					25					30			
Phe	Ala	Ala	Ala	Ser	Ser	Gly	Arg	Pro	Ser	Ser	Gln	Pro	Ser	Asp	Glu	
		35					40					45				
Val	Leu	Ser	Glu	Phe	Glu	Leu	Arg	Leu	Leu	Ser	Met	Phe	Gly	Leu	Lys	
	50					55					60					
Gln	Arg	Pro	Thr	Pro	Ser	Arg	Asp	Ala	Val	Val	Pro	Pro	Tyr	Met	Leu	
65					70					75					80	
Asp	Leu	Tyr	Arg	Arg	His	Ser	Gly	Gln	Pro	Gly	Ser	Pro	Ala	Pro	Asp	
				85					90					95		
His	Arg	Leu	Glu	Arg	Ala	Ala	Ser	Arg	Ala	Asn	Thr	Val	Arg	Ser	Phe	
			100					105					110			
His	His	Glu	Glu	Ser	Leu	Glu	Glu	Leu	Pro	Glu	Thr	Ser	Gly	Lys	Thr	
		115					120					125				
Thr	Arg	Arg	Phe	Phe	Phe	Asn	Leu	Ser	Ser	Ile	Pro	Thr	Glu	Glu	Phe	
	130					135					140					
Ile	Thr	Ser	Ala	Glu	Leu	Gln	Val	Phe	Arg	Glu	Gln	Met	Gln	Asp	Ala	
145					150					155					160	
Leu	Gly	Asn	Asn	Ser	Ser	Phe	His	His	Arg	Ile	Asn	Ile	Tyr	Glu	Ile	
				165					170					175		
Ile	Lys	Pro	Ala	Thr	Ala	Asn	Ser	Lys	Phe	Pro	Val	Thr	Arg	Leu	Leu	
			180					185					190			
Asp	Thr	Arg	Leu	Val	Asn	Gln	Asn	Ala	Ser	Arg	Trp	Glu	Thr	Phe	Asp	
		195					200					205				
Val	Thr	Pro	Ala	Val	Met	Arg	Trp	Thr	Ala	Gln	Gly	His	Ala	Asn	His	
	210					215					220					
Gly	Phe	Val	Val	Glu	Val	Ala	His	Leu	Glu	Glu	Lys	Gln	Gly	Val	Ser	
225					230					235					240	
Lys	Arg	His	Val	Arg	Ile	Ser	Arg	Ser	Leu	His	Gln	Asp	Glu	His	Ser	
				245					250					255		
Trp	Ser	Gln	Ile	Arg	Pro	Leu	Leu	Val	Thr	Phe	Gly	His	Asp	Gly	Lys	
			260					265					270			
Gly	His	Pro	Leu	His	Lys	Arg	Glu	Lys	Arg	Gln	Ala	Lys	His	Lys	Gln	

275

280

285

Arg Lys Arg Leu Lys Ser Ser Cys Lys Arg His Pro Leu Tyr Val Asp
 290 295 300
 Phe Ser Asp Val Gly Trp Asn Asp Trp Ile Val Ala Pro Pro Gly Tyr
 305 310 315 320
 His Ala Phe Tyr Cys His Gly Glu Cys Pro Phe Pro Leu Ala Asp His
 325 330 335
 Leu Asn Ser Thr Asn His Ala Ile Val Gln Thr Leu Val Asn Ser Val
 340 345 350
 Asn Ser Lys Ile Pro Lys Ala Cys Cys Val Pro Thr Glu Leu Ser Ala
 355 360 365
 Ile Ser Met Leu Tyr Leu Asp Glu Asn Glu Lys Val Val Leu Lys Asn
 370 375 380
 Tyr Gln Asp Met Val Val Glu Gly Cys Gly Cys Arg
 385 390 395

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1954 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 403..1626

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CTCTAGAGGG CAGAGGAGGA GGGAGGGAGG GAAGGAGCGC GGAGCCCCGGC CCGGAAGCTA	60
GGTGAGTGTG GCATCCGAGC TGAGGGACGC GAGCCTGAGA CGCCGCTGCT GCTCCGGCTG	120
AGTATCTAGC TTGTCTCCCC GATGGGATTC CCGTCCAAGC TATCTCGAGC CTGCAGCGCC	180
ACAGTCCCCG GCCCTCGCCC AGGTTCACCTG CAACCGTTCA GAGGTCCCCA GGAGCTGCTG	240
CTGGCGAGCC CGCTACTGCA GGGACCTATG GAGCCATTCC GTAGTGCCAT CCCGAGCAAC	300
GCACTGCTGC AGCTTCCCTG AGCCTTTCCA GCAAGTTTGT TCAAGATTGG CTGTCAAGAA	360
TCATGGACTG TTATTATATG CCTTGTTTTT TGTCAAGACA CC ATG ATT CCT GGT	414
Met Ile Pro Gly	
1	
AAC CGA ATG CTG ATG GTC GTT TTA TTA TGC CAA GTC CTG CTA GGA GGC	462
Asn Arg Met Leu Met Val Val Leu Leu Cys Gln Val Leu Leu Gly Gly	
5 10 15 20	

GCG Ala	AGC Ser	CAT His	GCT Ala	AGT Ser 25	TTG Leu	ATA Ile	CCT Pro	GAG Glu	ACG Thr 30	GGG Gly	AAG Lys	AAA Lys	AAA Lys	GTC Val 35	GCC Ala	510
GAG Glu	ATT Ile	CAG Gln 40	GGC Gly	CAC His	GCG Ala	GGA Gly	GGA Gly	CGC Arg 45	CGC Arg	TCA Ser	GGG Gly	CAG Gln 50	AGC Ser	CAT His	GAG Glu	558
CTC Leu	CTG Leu	CGG Arg 55	GAC Asp	TTC Phe	GAG Glu	GCG Ala	ACA Thr 60	CTT Leu	CTG Leu	CAG Gln	ATG Met	TTT Phe 65	GGG Gly	CTG Leu	CGC Arg	606
CGC Arg	CGC Arg 70	CCG Pro	CAG Gln	CCT Pro	AGC Ser	AAG Lys 75	AGT Ser	GCC Ala	GTC Val	ATT Ile	CCG Pro 80	GAC Asp	TAC Tyr	ATG Met	CGG Arg	654
GAT Asp 85	CTT Leu	TAC Tyr	CGG Arg	CTT Leu	CAG Gln 90	TCT Ser	GGG Gly	GAG Glu	GAG Glu	GAG Glu 95	GAA Glu	GAG Glu	CAG Gln	ATC Ile	CAC His 100	702
AGC Ser	ACT Thr	GGT Gly	CTT Leu 105	GAG Glu	TAT Tyr	CCT Pro	GAG Glu	CGC Arg	CCG Pro 110	GCC Ala	AGC Ser	CGG Arg	GCC Ala	AAC Asn 115	ACC Thr	750
GTG Val	AGG Arg	AGC Ser 120	TTC Phe	CAC His	CAC His	GAA Glu	GAA Glu	CAT His 125	CTG Leu	GAG Glu	AAC Asn	ATC Ile	CCA Pro 130	GGG Gly	ACC Thr	798
AGT Ser	GAA Glu 135	AAC Asn	TCT Ser	GCT Ala	TTT Phe	CGT Arg	TTC Phe 140	CTC Leu	TTT Phe	AAC Asn	CTC Leu	AGC Ser 145	AGC Ser	ATC Ile	CCT Pro	846
GAG Glu 150	AAC Asn	GAG Glu	GTG Val	ATC Ile	TCC Ser	TCT Ser 155	GCA Ala	GAG Glu	CTT Leu	CGG Arg 160	CTC Leu	TTC Phe	CGG Arg	GAG Glu	CAG Gln	894
GTG Val 165	GAC Asp	CAG Gln	GGC Gly	CCT Pro	GAT Asp 170	TGG Trp	GAA Glu	AGG Arg	GGC Gly	TTC Phe 175	CAC His	CGT Arg	ATA Ile	AAC Asn	ATT Ile 180	942
TAT Tyr	GAG Glu	GTT Val	ATG Met	AAG Lys 185	CCC Pro	CCA Pro	GCA Ala	GAA Glu	GTG Val 190	GTG Val	CCT Pro	GGG Gly	CAC His	CTC Leu 195	ATC Ile	990
ACA Thr	CGA Arg	CTA Leu 200	CTG Leu	GAC Asp	ACG Thr	AGA Arg	CTG Leu	GTC Val 205	CAC His	CAC His	AAT Asn	GTG Val	ACA Thr 210	CGG Arg	TGG Trp	1038
GAA Glu	ACT Thr 215	TTT Phe	GAT Asp	GTG Val	AGC Ser	CCT Pro	GCG Ala 220	GTC Val	CTT Leu	CGC Arg	TGG Trp	ACC Thr 225	CGG Arg	GAG Glu	AAG Lys	1086
CAG Gln 230	CCA Pro	AAC Asn	TAT Tyr	GGG Gly	CTA Leu	GCC Ala 235	ATT Ile	GAG Glu	GTG Val	ACT Thr	CAC His 240	CTC Leu	CAT His	CAG Gln	ACT Thr	1134
CGG Arg 245	ACC Thr	CAC His	CAG Gln	GGC Gly	CAG Gln 250	CAT His	GTC Val	AGG Arg	ATT Ile	AGC Ser 255	CGA Arg	TCG Ser	TTA Leu	CCT Pro	CAA Gln 260	1182

GGG	AGT	GGG	AAT	TGG	GCC	CAG	CTC	CGG	CCC	CTC	CTG	GTC	ACC	TTT	GGC	1230
Gly	Ser	Gly	Asn	Trp	Ala	Gln	Leu	Arg	Pro	Leu	Leu	Val	Thr	Phe	Gly	
			265						270					275		
CAT	GAT	GGC	CGG	GGC	CAT	GCC	TTG	ACC	CGA	CGC	CGG	AGG	GCC	AAG	CGT	1278
His	Asp	Gly	Arg	Gly	His	Ala	Leu	Thr	Arg	Arg	Arg	Arg	Ala	Lys	Arg	
			280					285					290			
AGC	CCT	AAG	CAT	CAC	TCA	CAG	CGG	GCC	AGG	AAG	AAG	AAT	AAG	AAC	TGC	1326
Ser	Pro	Lys	His	His	Ser	Gln	Arg	Ala	Arg	Lys	Lys	Asn	Lys	Asn	Cys	
		295					300					305				
CGG	CGC	CAC	TCG	CTC	TAT	GTG	GAC	TTC	AGC	GAT	GTG	GGC	TGG	AAT	GAC	1374
Arg	Arg	His	Ser	Leu	Tyr	Val	Asp	Phe	Ser	Asp	Val	Gly	Trp	Asn	Asp	
	310					315					320					
TGG	ATT	GTG	GCC	CCA	CCA	GGC	TAC	CAG	GCC	TTC	TAC	TGC	CAT	GGG	GAC	1422
Trp	Ile	Val	Ala	Pro	Pro	Gly	Tyr	Gln	Ala	Phe	Tyr	Cys	His	Gly	Asp	
325					330					335					340	
TGC	CCC	TTT	CCA	CTG	GCT	GAC	CAC	CTC	AAC	TCA	ACC	AAC	CAT	GCC	ATT	1470
Cys	Pro	Phe	Pro	Leu	Ala	Asp	His	Leu	Asn	Ser	Thr	Asn	His	Ala	Ile	
				345					350					355		
GTG	CAG	ACC	CTG	GTC	AAT	TCT	GTC	AAT	TCC	AGT	ATC	CCC	AAA	GCC	TGT	1518
Val	Gln	Thr	Leu	Val	Asn	Ser	Val	Asn	Ser	Ser	Ile	Pro	Lys	Ala	Cys	
			360					365					370			
TGT	GTG	CCC	ACT	GAA	CTG	AGT	GCC	ATC	TCC	ATG	CTG	TAC	CTG	GAT	GAG	1566
Cys	Val	Pro	Thr	Glu	Leu	Ser	Ala	Ile	Ser	Met	Leu	Tyr	Leu	Asp	Glu	
		375					380					385				
TAT	GAT	AAG	GTG	GTA	CTG	AAA	AAT	TAT	CAG	GAG	ATG	GTA	GTA	GAG	GGA	1614
Tyr	Asp	Lys	Val	Val	Leu	Lys	Asn	Tyr	Gln	Glu	Met	Val	Val	Glu	Gly	
	390					395					400					
TGT	GGG	TGC	CGC	TGAGATCAGG	CAGTCCTTGA	GGATAGACAG	ATATACACAC									1666
Cys	Gly	Cys	Arg													
405																
CACACACACA	CACCACATAC	ACCACACACA	CACGTTCCCA	TCCACTCACC	CACACACTAC											1726
ACAGACTGCT	TCCTTATAGC	TGGACTTTTA	TTTAAAAAAA	AAAAAATAAA	AATGGAAAAA											1786
ATCCCTAAAC	ATTCACCTTG	ACCTTATTTA	TGACTTTACG	TGCAAATGTT	TTGACCATAT											1846
TGATCATATA	TTTTGACAAA	ATATATTTAT	AACTACGTAT	TAAAAGAAAA	AAATAAAATG											1906
AGTCATTATT	TTAAAAAATA	AAAAAATACT	CTAGAGTCGA	CGGAATTC												1954

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 408 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met 1	Ile	Pro	Gly	Asn 5	Arg	Met	Leu	Met	Val 10	Val	Leu	Leu	Cys	Gln 15	Val
Leu	Leu	Gly	Gly 20	Ala	Ser	His	Ala	Ser 25	Leu	Ile	Pro	Glu	Thr 30	Gly	Lys
Lys	Lys 35	Val	Ala	Glu	Ile	Gln	Gly 40	His	Ala	Gly	Gly	Arg 45	Arg	Ser	Gly
Gln	Ser 50	His	Glu	Leu	Leu	Arg 55	Asp	Phe	Glu	Ala	Thr 60	Leu	Leu	Gln	Met
Phe 65	Gly	Leu	Arg	Arg	Arg 70	Pro	Gln	Pro	Ser	Lys 75	Ser	Ala	Val	Ile	Pro 80
Asp	Tyr	Met	Arg	Asp 85	Leu	Tyr	Arg	Leu	Gln 90	Ser	Gly	Glu	Glu	Glu 95	Glu
Glu	Gln	Ile	His 100	Ser	Thr	Gly	Leu	Glu 105	Tyr	Pro	Glu	Arg	Pro 110	Ala	Ser
Arg	Ala	Asn 115	Thr	Val	Arg	Ser	Phe 120	His	His	Glu	Glu	His 125	Leu	Glu	Asn
Ile 130	Pro	Gly	Thr	Ser	Glu	Asn 135	Ser	Ala	Phe	Arg	Phe 140	Leu	Phe	Asn	Leu
Ser 145	Ser	Ile	Pro	Glu	Asn 150	Glu	Val	Ile	Ser	Ser 155	Ala	Glu	Leu	Arg	Leu 160
Phe	Arg	Glu	Gln 165	Val	Asp	Gln	Gly	Pro	Asp 170	Trp	Glu	Arg	Gly	Phe 175	His
Arg	Ile	Asn 180	Ile	Tyr	Glu	Val	Met	Lys 185	Pro	Pro	Ala	Glu	Val 190	Val	Pro
Gly	His 195	Leu	Ile	Thr	Arg	Leu	Leu	Asp 200	Thr	Arg	Leu	Val 205	His	His	Asn
Val 210	Thr	Arg	Trp	Glu	Thr	Phe 215	Asp	Val	Ser	Pro	Ala 220	Val	Leu	Arg	Trp
Thr 225	Arg	Glu	Lys	Gln	Pro 230	Asn	Tyr	Gly	Leu	Ala 235	Ile	Glu	Val	Thr	His 240
Leu	His	Gln	Thr 245	Arg	Thr	His	Gln	Gly	Gln 250	His	Val	Arg	Ile	Ser 255	Arg
Ser	Leu	Pro	Gln 260	Gly	Ser	Gly	Asn	Trp 265	Ala	Gln	Leu	Arg	Pro 270	Leu	Leu
Val	Thr 275	Phe	Gly	His	Asp	Gly	Arg 280	Gly	His	Ala	Leu	Thr 285	Arg	Arg	Arg
Arg 290	Ala	Lys	Arg	Ser	Pro	Lys 295	His	His	Ser	Gln	Arg 300	Ala	Arg	Lys	Lys

Asn	Lys	Asn	Cys	Arg	Arg	His	Ser	Leu	Tyr	Val	Asp	Phe	Ser	Asp	Val
305					310					315					320
Gly	Trp	Asn	Asp	Trp	Ile	Val	Ala	Pro	Pro	Gly	Tyr	Gln	Ala	Phe	Tyr
				325					330					335	
Cys	His	Gly	Asp	Cys	Pro	Phe	Pro	Leu	Ala	Asp	His	Leu	Asn	Ser	Thr
			340					345					350		
Asn	His	Ala	Ile	Val	Gln	Thr	Leu	Val	Asn	Ser	Val	Asn	Ser	Ser	Ile
		355					360					365			
Pro	Lys	Ala	Cys	Cys	Val	Pro	Thr	Glu	Leu	Ser	Ala	Ile	Ser	Met	Leu
	370					375					380				
Tyr	Leu	Asp	Glu	Tyr	Asp	Lys	Val	Val	Leu	Lys	Asn	Tyr	Gln	Glu	Met
385					390					395					400
Val	Val	Glu	Gly	Cys	Gly	Cys	Arg								
				405											

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1448 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 97..1389

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GTGACCGAGC	GGCGCGGACG	GCCGCCTGCC	CCCTCTGCCA	CCTGGGGCGG	TGCGGGCCCG		60
GAGCCCGGAG	CCCGGGTAGC	GCGTAGAGCC	GGCGCG	ATG	CAC	GTG	CGC
				Met	His	Val	Arg
				1			5
CGA	GCT	GCG	GCG	CCG	CAC	AGC	TTC
Arg	Ala	Ala	Ala	Pro	His	Ser	Phe
			10				15
CTG	CTG	CGC	TCC	GCC	CTG	GCC	GAC
Leu	Leu	Arg	Ser	Ala	Leu	Ala	Asp
		25				30	
TCG	AGC	TTC	ATC	CAC	CGG	CGC	CTC
Ser	Ser	Phe	Ile	His	Arg	Arg	Leu
		40				45	
CAG	CGC	GAG	ATC	CTC	TCC	ATT	TTG
Gln	Arg	Glu	Ile	Leu	Ser	Ile	Leu
		55				60	
							65
							70

CAC His	CTC Leu	CAG Gln	GGC Gly	AAG Lys 75	CAC His	AAC Asn	TCG Ser	GCA Ala	CCC Pro 80	ATG Met	TTC Phe	ATG Met	CTG Leu	GAC Asp 85	CTG Leu	354
TAC Tyr	AAC Asn	GCC Ala	ATG Met 90	GCG Ala	GTG Val	GAG Glu	GAG Glu	GGC Gly 95	GGC Gly	GGG Gly	CCC Pro	GGC Gly	GGC Gly 100	CAG Gln	GGC Gly	402
TTC Phe	TCC Ser	TAC Tyr 105	CCC Pro	TAC Tyr	AAG Lys	GCC Ala	GTC Val 110	TTC Phe	AGT Ser	ACC Thr	CAG Gln	GGC Gly 115	CCC Pro	CCT Pro	CTG Leu	450
GCC Ala	AGC Ser 120	CTG Leu	CAA Gln	GAT Asp	AGC Ser	CAT His 125	TTC Phe	CTC Leu	ACC Thr	GAC Asp	GCC Ala 130	GAC Asp	ATG Met	GTC Val	ATG Met	498
AGC Ser 135	TTC Phe	GTC Val	AAC Asn	CTC Leu	GTG Val 140	GAA Glu	CAT His	GAC Asp	AAG Lys	GAA Glu 145	TTC Phe	TTC Phe	CAC His	CCA Pro	CGC Arg 150	546
TAC Tyr	CAC His	CAT His	CGA Arg	GAG Glu 155	TTC Phe	CGG Arg	TTT Phe	GAT Asp	CTT Leu 160	TCC Ser	AAG Lys	ATC Ile	CCA Pro	GAA Glu 165	GGG Gly	594
GAA Glu	GCT Ala	GTC Val	ACG Thr 170	GCA Ala	GCC Ala	GAA Glu	TTC Phe	CGG Arg 175	ATC Ile	TAC Tyr	AAG Lys	GAC Asp	TAC Tyr 180	ATC Ile	CGG Arg	642
GAA Glu	CGC Arg	TTC Phe 185	GAC Asp	AAT Asn	GAG Glu	ACG Thr	TTC Phe 190	CGG Arg	ATC Ile	AGC Ser	GTT Val	TAT Tyr 195	CAG Gln	GTG Val	CTC Leu	690
CAG Gln 200	GAG Glu	CAC His	TTG Leu	GGC Gly	AGG Arg	GAA Glu 205	TCG Ser	GAT Asp	CTC Leu	TTC Phe	CTG Leu 210	CTC Leu	GAC Asp	AGC Ser	CGT Arg	738
ACC Thr 215	CTC Leu	TGG Trp	GCC Ala	TCG Ser	GAG Glu 220	GAG Glu	GGC Gly	TGG Trp	CTG Leu	GTG Val 225	TTT Phe	GAC Asp	ATC Ile	ACA Thr	GCC Ala 230	786
ACC Thr	AGC Ser	AAC Asn	CAC His	TGG Trp 235	GTG Val	GTC Val	AAT Asn	CCG Pro	CGG Arg 240	CAC His	AAC Asn	CTG Leu	GGC Gly 245	CTG Leu	CAG Gln	834
CTC Leu	TCG Ser	GTG Val	GAG Glu 250	ACG Thr	CTG Leu	GAT Asp	GGG Gly	CAG Gln 255	AGC Ser	ATC Ile	AAC Asn	CCC Pro	AAG Lys 260	TTG Leu	GCG Ala	882
GGC Gly	CTG Leu	ATT Ile 265	GGG Gly	CGG Arg	CAC His	GGG Gly	CCC Pro 270	CAG Gln	AAC Asn	AAG Lys	CAG Gln	CCC Pro 275	TTC Phe	ATG Met	GTG Val	930
GCT Ala	TTC Phe 280	TTC Phe	AAG Lys	GCC Ala	ACG Thr	GAG Glu 285	GTC Val	CAC His	TTC Phe	CGC Arg	AGC Ser 290	ATC Ile	CGG Arg	TCC Ser	ACG Thr	978
GGG Gly 295	AGC Ser	AAA Lys	CAG Gln	CGC Arg	AGC Ser 300	CAG Gln	AAC Asn	CGC Arg	TCC Ser	AAG Lys 305	ACG Thr	CCC Pro	AAG Lys	AAC Asn	CAG Gln 310	1026

GAA	GCC	CTG	CGG	ATG	GCC	AAC	GTG	GCA	GAG	AAC	AGC	AGC	AGC	GAC	CAG	1074
Glu	Ala	Leu	Arg	Met	Ala	Asn	Val	Ala	Glu	Asn	Ser	Ser	Ser	Asp	Gln	
				315					320					325		
AGG	CAG	GCC	TGT	AAG	AAG	CAC	GAG	CTG	TAT	GTC	AGC	TTC	CGA	GAC	CTG	1122
Arg	Gln	Ala	Cys	Lys	Lys	His	Glu	Leu	Tyr	Val	Ser	Phe	Arg	Asp	Leu	
			330					335					340			
GGC	TGG	CAG	GAC	TGG	ATC	ATC	GCG	CCT	GAA	GGC	TAC	GCC	GCC	TAC	TAC	1170
Gly	Trp	Gln	Asp	Trp	Ile	Ile	Ala	Pro	Glu	Gly	Tyr	Ala	Ala	Tyr	Tyr	
		345					350					355				
TGT	GAG	GGG	GAG	TGT	GCC	TTC	CCT	CTG	AAC	TCC	TAC	ATG	AAC	GCC	ACC	1218
Cys	Glu	Gly	Glu	Cys	Ala	Phe	Pro	Leu	Asn	Ser	Tyr	Met	Asn	Ala	Thr	
	360					365					370					
AAC	CAC	GCC	ATC	GTG	CAG	ACG	CTG	GTC	CAC	TTC	ATC	AAC	CCG	GAA	ACG	1266
Asn	His	Ala	Ile	Val	Gln	Thr	Leu	Val	His	Phe	Ile	Asn	Pro	Glu	Thr	
					380					385					390	
GTG	CCC	AAG	CCC	TGC	TGT	GCG	CCC	ACG	CAG	CTC	AAT	GCC	ATC	TCC	GTC	1314
Val	Pro	Lys	Pro	Cys	Cys	Ala	Pro	Thr	Gln	Leu	Asn	Ala	Ile	Ser	Val	
				395					400					405		
CTC	TAC	TTC	GAT	GAC	AGC	TCC	AAC	GTC	ATC	CTG	AAG	AAA	TAC	AGA	AAC	1362
Leu	Tyr	Phe	Asp	Asp	Ser	Ser	Asn	Val	Ile	Leu	Lys	Lys	Tyr	Arg	Asn	
			410					415					420			
ATG	GTG	GTC	CGG	GCC	TGT	GGC	TGC	CAC	TAGCTCCTCC GAGAATTCAG						1409	
Met	Val	Val	Arg	Ala	Cys	Gly	Cys	His								
		425				430										
ACCCTTTGGG GCCAAGTTTT TCTGGATCCT CCATTGCTC																1448

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 431 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met	His	Val	Arg	Ser	Leu	Arg	Ala	Ala	Ala	Pro	His	Ser	Phe	Val	Ala
1				5					10					15	
Leu	Trp	Ala	Pro	Leu	Phe	Leu	Leu	Arg	Ser	Ala	Leu	Ala	Asp	Phe	Ser
			20					25					30		
Leu	Asp	Asn	Glu	Val	His	Ser	Ser	Phe	Ile	His	Arg	Arg	Leu	Arg	Ser
		35					40					45			
Gln	Glu	Arg	Arg	Glu	Met	Gln	Arg	Glu	Ile	Leu	Ser	Ile	Leu	Gly	Leu
	50					55					60				
Pro	His	Arg	Pro	Arg	Pro	His	Leu	Gln	Gly	Lys	His	Asn	Ser	Ala	Pro

65					70					75					80				
Met	Phe	Met	Leu	Asp 85	Leu	Tyr	Asn	Ala	Met 90	Ala	Val	Glu	Glu	Gly 95	Gly				
Gly	Pro	Gly	Gly 100	Gln	Gly	Phe	Ser	Tyr 105	Pro	Tyr	Lys	Ala	Val 110	Phe	Ser				
Thr	Gln	Gly 115	Pro	Pro	Leu	Ala	Ser 120	Leu	Gln	Asp	Ser	His 125	Phe	Leu	Thr				
Asp	Ala 130	Asp	Met	Val	Met	Ser 135	Phe	Val	Asn	Leu	Val 140	Glu	His	Asp	Lys				
Glu 145	Phe	Phe	His	Pro	Arg 150	Tyr	His	His	Arg	Glu 155	Phe	Arg	Phe	Asp	Leu 160				
Ser	Lys	Ile	Pro	Glu 165	Gly	Glu	Ala	Val	Thr 170	Ala	Ala	Glu	Phe	Arg 175	Ile				
Tyr	Lys	Asp	Tyr 180	Ile	Arg	Glu	Arg	Phe 185	Asp	Asn	Glu	Thr	Phe 190	Arg	Ile				
Ser	Val	Tyr 195	Gln	Val	Leu	Gln	Glu 200	His	Leu	Gly	Arg	Glu 205	Ser	Asp	Leu				
Phe	Leu 210	Leu	Asp	Ser	Arg	Thr 215	Leu	Trp	Ala	Ser	Glu 220	Glu	Gly	Trp	Leu				
Val 225	Phe	Asp	Ile	Thr	Ala 230	Thr	Ser	Asn	His	Trp 235	Val	Val	Asn	Pro	Arg 240				
His	Asn	Leu	Gly	Leu 245	Gln	Leu	Ser	Val	Glu 250	Thr	Leu	Asp	Gly	Gln 255	Ser				
Ile	Asn	Pro	Lys 260	Leu	Ala	Gly	Leu	Ile 265	Gly	Arg	His	Gly	Pro 270	Gln	Asn				
Lys	Gln	Pro	Phe	Met	Val	Ala	Phe 280	Phe	Lys	Ala	Thr	Glu 285	Val	His	Phe				
Arg	Ser 290	Ile	Arg	Ser	Thr	Gly 295	Ser	Lys	Gln	Arg	Ser 300	Gln	Asn	Arg	Ser				
Lys 305	Thr	Pro	Lys	Asn	Gln 310	Glu	Ala	Leu	Arg	Met 315	Ala	Asn	Val	Ala	Glu 320				
Asn	Ser	Ser	Ser	Asp 325	Gln	Arg	Gln	Ala	Cys 330	Lys	Lys	His	Glu	Leu 335	Tyr				
Val	Ser	Phe	Arg	Asp	Leu	Gly	Trp	Gln 345	Asp	Trp	Ile	Ile	Ala 350	Pro	Glu				
Gly	Tyr	Ala	Ala	Tyr	Tyr	Cys	Glu 360	Gly	Glu	Cys	Ala	Phe 365	Pro	Leu	Asn				
Ser	Tyr 370	Met	Asn	Ala	Thr	Asn 375	His	Ala	Ile	Val	Gln 380	Thr	Leu	Val	His				
Val	Ile	Asn	Pro	Glu	Thr	Val	Pro	Lys	Pro	Cys	Cys	Ala	Pro	Thr	Gln				

Leu Asn Ala Ile Ser Val Leu Tyr Phe Asp Asp Ser Ser Asn Val Ile

405410415

Leu Lys Lys Tyr Arg Asn Met Val Val Arg Ala Cys Gly Cys His

420425430

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2923 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo sapiens
 - (F) TISSUE TYPE: Human placenta
- (vii) IMMEDIATE SOURCE:
 - (A) LIBRARY: Stratagene catalog #936203 Human placenta cDNA library
 - (B) CLONE: BMP6C35
- (viii) POSITION IN GENOME:
 - (C) UNITS: bp
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 160..1701
- (ix) FEATURE:
 - (A) NAME/KEY: mat_peptide
 - (B) LOCATION: 1282..1698
- (ix) FEATURE:
 - (A) NAME/KEY: mRNA
 - (B) LOCATION: 1..2923

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CGACCATGAG AGATAAGGAC TGAGGGCCAG GAAGGGGAAG CGAGCCCGCC GAGAGGTGGC60

GGGGACTGCT CACGCCAAGG GCCACAGCGG CCGCGCTCCG GCCTCGCTCC GCCGCTCCAC120

GCCTCGCGGG ATCCGCGGGG GCAGCCCGGC CGGGCGGGG ATG CCG GGG CTG GGG174

Met Pro Gly Leu Gly-374-370

CGG AGG GCG CAG TGG CTG TGC TGG TGG TGG GGG CTG CTG TGC AGC TGC222

Arg Arg Ala Gln Trp Leu Cys Trp Trp Trp Gly Leu Leu Cys Ser Cys-365-360-355

TGC GGG CCC CCG CCG CTG CGG CCG CCC TTG CCC GCT GCC GCG GCC GCC270

Cys Gly Pro Pro Pro Leu Ala Pro Pro Leu Pro Ala Ala Ala Ala	-350	-345	-340	
GCC GCC GGG GGG CAG CTG CTG GGG GAC GGC GGG AGC CCC GGC CGC ACG				318
Ala Ala Gly Gly Gln Leu Leu Gly Asp Gly Gly Ser Pro Gly Arg Thr	-335	-330	-325	
GAG CAG CCG CCG CCG TCG CCG CAG TCC TCC TCG GGC TTC CTG TAC CGG				366
Glu Gln Pro Pro Pro Ser Pro Gln Ser Ser Ser Gly Phe Leu Tyr Arg	-320	-315	-310	
CGG CTC AAG ACG CAG GAG AAG CGG GAG ATG CAG AAG GAG ATC TTG TCG				414
Arg Leu Lys Thr Gln Glu Lys Arg Glu Met Gln Lys Glu Ile Leu Ser	-305	-300	-295	-290
GTG CTG GGG CTC CCG CAC CGG CCC CGG CCC CTG CAC GGC CTC CAA CAG				462
Val Leu Gly Leu Pro His Arg Pro Arg Pro Leu His Gly Leu Gln Gln	-285		-280	-275
CCG CAG CCC CCG GCG CTC CGG CAG CAG GAG GAG CAG CAG CAG CAG CAG				510
Pro Gln Pro Pro Ala Leu Arg Gln Gln Glu Glu Gln Gln Gln Gln Gln	-270	-265	-260	
CAG CTG CCT CGC GGA GAG CCC CCT CCC GGG CGA CTG AAG TCC GCG CCC				558
Gln Leu Pro Arg Gly Glu Pro Pro Pro Gly Arg Leu Lys Ser Ala Pro	-255	-250	-245	
CTC TTC ATG CTG GAT CTG TAC AAC GCC CTG TCC GCC GAC AAC GAC GAG				606
Leu Phe Met Leu Asp Leu Tyr Asn Ala Leu Ser Ala Asp Asn Asp Glu	-240	-235	-230	
GAC GGG GCG TCG GAG GGG GAG AGG CAG CAG TCC TGG CCC CAC GAA GCA				654
Asp Gly Ala Ser Glu Gly Glu Arg Gln Gln Ser Trp Pro His Glu Ala	-225	-220	-215	-210
GCC AGC TCG TCC CAG CGT CGG CAG CCG CCC CCG GGC GCC GCG CAC CCG				702
Ala Ser Ser Ser Gln Arg Arg Gln Pro Pro Pro Gly Ala Ala His Pro	-205	-200	-195	
CTC AAC CGC AAG AGC CTT CTG GCC CCC GGA TCT GGC AGC GGC GGC GCG				750
Leu Asn Arg Lys Ser Leu Leu Ala Pro Gly Ser Gly Ser Gly Gly Ala	-190	-185	-180	
TCC CCA CTG ACC AGC GCG CAG GAC AGC GCC TTC CTC AAC GAC GCG GAC				798
Ser Pro Leu Thr Ser Ala Gln Asp Ser Ala Phe Leu Asn Asp Ala Asp	-175	-170	-165	
ATG GTC ATG AGC TTT GTG AAC CTG GTG GAG TAC GAC AAG GAG TTC TCC				846
Met Val Met Ser Phe Val Asn Leu Val Glu Tyr Asp Lys Glu Phe Ser	-160	-155	-150	
CCT CGT CAG CGA CAC CAC AAA GAG TTC AAG TTC AAC TTA TCC CAG ATT				894
Pro Arg Gln Arg His His Lys Glu Phe Lys Phe Asn Leu Ser Gln Ile	-145	-140	-135	-130
CCT GAG GGT GAG GTG GTG ACG GCT GCA GAA TTC CGC ATC TAC AAG GAC				942
Pro Glu Gly Glu Val Val Thr Ala Ala Glu Phe Arg Ile Tyr Lys Asp	-125	-120	-115	
TGT GTT ATG GGG AGT TTT AAA AAC CAA ACT TTT CTT ATC AGC ATT TAT				990

Cys	Val	Met	Gly	Ser	Phe	Leu	Asn	Gln	Thr	Phe	Leu	Ile	Ser	Ile	Tyr	
			-110					-105					-100			
CAA	GTC	TTA	CAG	GAG	CAT	CAG	CAC	AGA	GAC	TCT	GAC	CTG	TTT	TTG	TTG	1038
Gln	Val	Leu	Gln	Glu	His	Gln	His	Arg	Asp	Ser	Asp	Leu	Phe	Leu	Leu	
		-95					-90					-85				
GAC	ACC	CGT	GTA	GTA	TGG	GCC	TCA	GAA	GAA	GGC	TGG	CTG	GAA	TTT	GAC	1086
Asp	Thr	Arg	Val	Val	Trp	Ala	Ser	Glu	Glu	Gly	Trp	Leu	Glu	Phe	Asp	
	-80					-75					-70					
ATC	ACG	GCC	ACT	AGC	AAT	CTG	TGG	GTT	GTG	ACT	CCA	CAG	CAT	AAC	ATG	1134
Ile	Thr	Ala	Thr	Ser	Asn	Leu	Trp	Val	Val	Thr	Pro	Gln	His	Asn	Met	
	-65				-60					-55					-50	
GGG	CTT	CAG	CTG	AGC	GTG	GTG	ACA	AGG	GAT	GGA	GTC	CAC	GTC	CAC	CCC	1182
Gly	Leu	Gln	Leu	Ser	Val	Val	Thr	Arg	Asp	Gly	Val	His	Val	His	Pro	
				-45					-40					-35		
CGA	GCC	GCA	GGC	CTG	GTG	GGC	AGA	GAC	GGC	CCT	TAC	GAT	AAG	CAG	CCC	1230
Arg	Ala	Ala	Gly	Leu	Val	Gly	Arg	Asp	Gly	Pro	Tyr	Asp	Lys	Gln	Pro	
			-30					-25					-20			
TTC	ATG	GTG	GCT	TTC	TTC	AAA	GTG	AGT	GAG	GTC	CAC	GTG	CGC	ACC	ACC	1278
Phe	Met	Val	Ala	Phe	Phe	Lys	Val	Ser	Glu	Val	His	Val	Arg	Thr	Thr	
	-15						-10					-5				
AGG	TCA	GCC	TCC	AGC	CGG	CGC	CGA	CAA	CAG	AGT	CGT	AAT	CGC	TCT	ACC	1326
Arg	Ser	Ala	Ser	Ser	Arg	Arg	Arg	Gln	Gln	Ser	Arg	Asn	Arg	Ser	Thr	
	1				5					10					15	
CAG	TCC	CAG	GAC	GTG	GCG	CGG	GTC	TCC	AGT	GCT	TCA	GAT	TAC	AAC	AGC	1374
Gln	Ser	Gln	Asp	Val	Ala	Arg	Val	Ser	Ser	Ala	Ser	Asp	Tyr	Asn	Ser	
				20				25						30		
AGT	GAA	TTG	AAA	ACA	GCC	TGC	AGG	AAG	CAT	GAG	CTG	TAT	GTG	AGT	TTC	1422
Ser	Glu	Leu	Lys	Thr	Ala	Cys	Arg	Lys	His	Glu	Leu	Tyr	Val	Ser	Phe	
			35					40					45			
CAA	GAC	CTG	GGA	TGG	CAG	GAC	TGG	ATC	ATT	GCA	CCC	AAG	GGC	TAT	GCT	1470
Gln	Asp	Leu	Gly	Trp	Gln	Asp	Trp	Ile	Ile	Ala	Pro	Lys	Gly	Tyr	Ala	
		50					55					60				
GCC	AAT	TAC	TGT	GAT	GGA	GAA	TGC	TCC	TTC	CCA	CTC	AAC	GCA	CAC	ATG	1518
Ala	Asn	Tyr	Cys	Asp	Gly	Glu	Cys	Ser	Phe	Pro	Leu	Asn	Ala	His	Met	
	65					70					75					
AAT	GCA	ACC	AAC	CAC	GCG	ATT	GTG	CAG	ACC	TTG	GTT	CAC	CTT	ATG	AAC	1566
Asn	Ala	Thr	Asn	His	Ala	Ile	Val	Gln	Thr	Leu	Val	His	Leu	Met	Asn	
	80				85					90					95	
CCC	GAG	TAT	GTC	CCC	AAA	CCG	TGC	TGT	GCG	CCA	ACT	AAG	CTA	AAT	GCC	1614
Pro	Glu	Tyr	Val	Pro	Lys	Pro	Cys	Cys	Ala	Pro	Thr	Lys	Leu	Asn	Ala	
				100					105					110		
ATC	TCG	GTT	CTT	TAC	TTT	GAT	GAC	AAC	TCC	AAT	GTC	ATT	CTG	AAA	AAA	1662
Ile	Ser	Val	Leu	Tyr	Phe	Asp	Asp	Asn	Ser	Asn	Val	Ile	Leu	Lys	Lys	
			115					120					125			
TAC	AGG	AAT	ATG	GTT	GTA	AGA	GCT	TGT	GGA	TGC	CAC	TAACTCGAAA				1708

Tyr Arg Asn Met Val Val Ala Cys Gly Cys His
130 135 140

CCAGATGCTG	GGGACACACA	TTCTGCCTTG	GATTCCTAGA	TTACATCTGC	CTTAAAAAAA	1768
CACGGAAGCA	CAGTTGGAGG	TGGGACGATG	AGACTTTGAA	ACTATCTCAT	GCCAGTGCCT	1828
TATTACCCAG	GAAGATTTTA	AAGGACCTCA	TTAATAATTT	GCTCACTTGG	TAAATGACGT	1888
GAGTAGTTGT	TGGTCTGTAG	CAAGCTGAGT	TTGGATGTCT	GTAGCATAAG	GTCTGGTAAC	1948
TGCAGAAACA	TAACCGTGAA	GCTCTTCCTA	CCCTCCTCCC	CCAAAAACCC	ACCAAAATTA	2008
GTTTTAGCTG	TAGATCAAGC	TATTTGGGGT	GTTTGTTAGT	AAATAGGGAA	AATAATCTCA	2068
AAGGAGTTAA	ATGTATTCTT	GGCTAAAGGA	TCAGCTGGTT	CAGTACTGTC	TATCAAAGGT	2128
AGATTTTACA	GAGAACAGAA	ATCGGGGAAG	TGGGGGGAAC	GCCTCTGTTC	AGTTCATTCC	2188
CAGAAGTCCA	CAGGACGCAC	AGCCCAGGCC	ACAGCCAGGG	CTCCACGGGG	CGCCCTTGTC	2248
TCAGTCATTG	CTGTTGTATG	TTCGTGCTGG	AGTTTTGTTG	GTGTGAAAAT	ACACTTATTT	2308
CAGCCAAAAC	ATACCATTTC	TACACCTCAA	TCCTCCATTT	GCTGTACTCT	TTGCTAGTAC	2368
CAAAGTAGA	CTGATTACAC	TGAGGTGAGG	CTACAAGGGG	TGTGTAACCG	TGTAACACGT	2428
GAAGGCAGTG	CTCACCTCTT	CTTTACCAGA	ACGGTTCTTT	GACCAGCACA	TTAACTTCTG	2488
GACTGCCGGC	TCTAGTACCT	TTTCAGTAAA	GTGGTTCTCT	GCCTTTTTTAC	TATACAGCAT	2548
ACCACGCCAC	AGGGTTAGAA	CCAACGAAGA	AAATAAAATG	AGGGTGCCCA	GCTTATAAGA	2608
ATGGTGTTAG	GGGGATGAGC	ATGCTGTTTA	TGAACGGAAA	TCATGATTTC	CCTGTAGAAA	2668
GTGAGGCTCA	GATTAAATTT	TAGAATATTT	TCTAAATGTC	TTTTTCACAA	TCATGTGACT	2728
GGGAAGGCAA	TTTCATACTA	AACTGATTAA	ATAATACATT	TATAATCTAC	AACTGTTTGC	2788
ACTTACAGCT	TTTTTTGTAA	ATATAAACTA	TAATTTATTG	TCTATTTTAT	ATCTGTTTTG	2848
CTGTGGCGTT	GGGGGGGGGG	CCGGGCTTTT	GGGGGGGGGG	GTTTGTTTGG	GGGGTGTCGT	2908
GGTGTGGGCG	GGCGG					2923

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 513 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Pro Gly Leu Gly Arg Arg Ala Gln Trp Leu Cys Trp Trp Trp Gly
-374 -370 -365 -360

Leu Leu Cys Ser Cys Cys Gly Pro Pro Pro Leu Arg Pro Pro Leu Pro
 -355 -350 -345

Ala Ala Ala Ala Ala Ala Gly Gly Gln Leu Leu Gly Asp Gly Gly
 -340 -335 -330

Ser Pro Gly Arg Thr Glu Gln Pro Pro Pro Ser Pro Gln Ser Ser Ser
 -325 -320 -315

Gly Phe Leu Tyr Arg Arg Leu Lys Thr Gln Glu Lys Arg Glu Met Gln
 -310 -305 -300 -295

Lys Glu Ile Leu Ser Val Leu Gly Leu Pro His Arg Pro Arg Pro Leu
 -290 -285 -280

His Gly Leu Gln Gln Pro Gln Pro Pro Ala Leu Arg Gln Gln Glu Glu
 -275 -270 -265

Gln Gln Gln Gln Gln Gln Leu Pro Arg Gly Glu Pro Pro Pro Gly Arg
 -260 -255 -250

Leu Lys Ser Ala Pro Leu Phe Met Leu Asp Leu Tyr Asn Ala Leu Ser
 -245 -240 -235

Ala Asp Asn Asp Glu Asp Gly Ala Ser Glu Gly Glu Arg Gln Gln Ser
 -230 -225 -220 -215

Trp Pro His Glu Ala Ala Ser Ser Ser Gln Arg Arg Gln Pro Pro Pro
 -210 -205 -200

Gly Ala Ala His Pro Leu Asn Arg Lys Ser Leu Leu Ala Pro Gly Ser
 -195 -190 -185

Gly Ser Gly Gly Ala Ser Pro Leu Thr Ser Ala Gln Asp Ser Ala Phe
 -180 -175 -170

Leu Asn Asp Ala Asp Met Val Met Ser Phe Val Asn Leu Val Glu Tyr
 -165 -160 -155

Asp Lys Glu Phe Ser Pro Arg Gln Arg His His Lys Glu Phe Lys Phe
 -150 -145 -140 -135

Asn Leu Ser Gln Ile Pro Glu Gly Glu Val Val Thr Ala Ala Glu Phe
 -130 -125 -120

Arg Ile Tyr Lys Asp Cys Val Met Gly Ser Phe Lys Asn Gln Thr Phe
 -115 -110 -105

Leu Ile Ser Ile Tyr Gln Val Leu Gln Glu His Gln His Arg Asp Ser
 -100 -95 -90

Asp Leu Phe Leu Leu Asp Thr Arg Val Val Trp Ala Ser Glu Glu Gly
 -85 -80 -75

Trp Leu Glu Phe Asp Ile Thr Ala Thr Ser Asn Leu Trp Val Val Thr
 -70 -65 -60 -55

Pro Gln His Asn Met Gly Leu Gln Leu Ser Val Val Thr Arg Asp Gly
 -50 -45 -40

Val His Val His Pro Arg Ala Ala Gly Leu Val Gly Arg Asp Gly Pro
 -35 -30 -25
 Tyr Asp Lys Gln Pro Phe Met Val Ala Phe Phe Lys Val Ser Glu Val
 -20 -15 -10
 His Val Arg Thr Thr Arg Ser Ala Ser Ser Arg Arg Arg Gln Gln Ser
 -5 1 5 10
 Arg Asn Arg Ser Thr Gln Ser Gln Asp Val Ala Arg Val Ser Ser Ala
 15 20 25
 Ser Asp Tyr Asn Ser Ser Glu Leu Lys Thr Ala Cys Arg Lys His Glu
 30 35 40
 Leu Tyr Val Ser Phe Gln Asp Leu Gly Trp Gln Asp Trp Ile Ile Ala
 45 50 55
 Pro Lys Gly Tyr Ala Ala Asn Tyr Cys Asp Gly Glu Cys Ser Phe Pro
 60 65 70
 Leu Asn Ala His Met Asn Ala Thr Asn His Ala Ile Val Gln Thr Leu
 75 80 85 90
 Val His Leu Met Asn Pro Glu Tyr Val Pro Lys Pro Cys Cys Ala Pro
 95 100 105
 Thr Lys Leu Asn Ala Ile Ser Val Leu Tyr Phe Asp Asp Asn Ser Asn
 110 115 120
 Val Ile Leu Lys Lys Tyr Arg Asn Met Val Val Arg Ala Cys Gly Cys
 125 130 135
 His

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2153 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo sapiens
- (H) CELL LINE: U2-OS osteosarcoma

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: U2-OS human osteosarcoma cDNA library
- (B) CLONE: U2-16

(viii) POSITION IN GENOME:

- (C) UNITS: bp

(ix) FEATURE:

- (A) NAME/KEY: CDS

(B) LOCATION: 69..2063

(ix) FEATURE:

(A) NAME/KEY: mat_peptide

(B) LOCATION: 1647..2060

. (ix) FEATURE:

(A) NAME/KEY: mRNA

(B) LOCATION: 1..2153

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CTGGTATATT TGTGCCTGCT GGAGGTGGAA TTAACAGTAA GAAGGAGAAA GGGATTGAAT	60
GGACTTACAG GAAGGATTTC AAGTAAATTC AGGGAAACAC ATTTACTTGA ATAGTACAAC	120
CTAGAGTATT ATTTTACACT AAGACGACAC AAAAGATGTT AAAGTTATCA CCAAGCTGCC	180
GGACAGATAT ATATTCCAAC ACCAAGGTGC AGATCAGCAT AGATCTGTGA TTCAGAAATC	240
AGGATTTGTT TTGGAAAGAG CTCAAGGGTT GAGAAGAACT CAAAAGCAAG TGAAGATTAC	300
TTTGGGAACT ACAGTTTATC AGAAGATCAA CTTTGTCTAA TTCAAATACC AAAGGCCTGA	360
TTATCATAAA TTCATATAGG AATGCATAGG TCATCTGATC AAATAATATT AGCCGTCTTC	420
TGCTACATCA ATGCAGCAAA AACTCTTAAC AACTGTGGAT AATTGGAAAT CTGAGTTTCA	480
GCTTTCTTAG AAATAACTAC TCTTGACATA TTCCAAAATA TTTAAAATAG GACAGGAAAA	540
TCGGTGAGGA TGTTGTGCTC AGAAATGTCA CTGTCATGAA AAATAGGTAA ATTTGTTTTT	600
TCAGCTACTG GGAAACTGTA CCTCCTAGAA CCTTAGGTTT TTTTTTTTTT AAGAGGACAA	660
GAAGGACTAA AAATATCAAC TTTTGCTTTT GGACAAAA ATG CAT CTG ACT GTA	713
Met His Leu Thr Val	
-316-315	
TTT TTA CTT AAG GGT ATT GTG GGT TTC CTC TGG AGC TGC TGG GTT CTA	761
Phe Leu Leu Lys Gly Ile Val Gly Phe Leu Trp Ser Cys Trp Val Leu	
-310 -305 -300	
GTG GGT TAT GCA AAA GGA GGT TTG GGA GAC AAT CAT GTT CAC TCC AGT	809
Val Gly Tyr Ala Lys Gly Gly Leu Gly Asp Asn His Val His Ser Ser	
-295 -290 -285 -280	
TTT ATT TAT AGA AGA CTA CGG AAC CAC GAA AGA CGG GAA ATA CAA AGG	857
Phe Ile Tyr Arg Arg Leu Arg Asn His Glu Arg Arg Glu Ile Gln Arg	
-275 -270 -265	
GAA ATT CTC TCT ATC TTG GGT TTG CCT CAC AGA CCC AGA CCA TTT TCA	905
Glu Ile Leu Ser Ile Leu Gly Leu Pro His Arg Pro Arg Pro Phe Ser	
-260 -255 -250	
CCT GGA AAA ATG ACC AAT CAA GCG TCC TCT GCA CCT CTC TTT ATG CTG	953
Pro Gly Lys Met Thr Asn Gln Ala Ser Ser Ala Pro Leu Phe Met Leu	
-245 -240 -235	
GAT CTC TAC AAT GCC GAA GAA AAT CCT GAA GAG TCG GAG TAC TCA GTA	1001

Asp	Leu	Tyr	Asn	Ala	Glu	Glu	Asn	Pro	Glu	Glu	Ser	Glu	Tyr	Ser	Val	
-230						-225					-220					
AGG	GCA	TCC	TTG	GCA	GAA	GAG	ACC	AGA	GGG	GCA	AGA	AAG	GGA	TAC	CCA	1049
Arg	Ala	Ser	Leu	Ala	Glu	Glu	Thr	Arg	Gly	Ala	Arg	Lys	Gly	Tyr	Pro	
-215					-210				-205						-200	
GCC	TCT	CCC	AAT	GGG	TAT	CCT	CGT	CGC	ATA	CAG	TTA	TCT	CGG	ACG	ACT	1097
Ala	Ser	Pro	Asn	Gly	Tyr	Pro	Arg	Arg	Ile	Gln	Leu	Ser	Arg	Thr	Thr	
				-195					-190					-185		
CCT	CTG	ACC	ACC	CAG	AGT	CCT	CCT	CTA	GCC	AGC	CTC	CAT	GAT	ACC	AAC	1145
Pro	Leu	Thr	Thr	Gln	Ser	Pro	Pro	Leu	Ala	Ser	Leu	His	Asp	Thr	Asn	
			-180					-175					-170			
TTT	CTG	AAT	GAT	GCT	GAC	ATG	GTC	ATG	AGC	TTT	GTC	AAC	TTA	GTT	GAA	1193
Phe	Leu	Asn	Asp	Ala	Asp	Met	Val	Met	Ser	Phe	Val	Asn	Leu	Val	Glu	
	-165					-160						-155				
AGA	GAC	AAG	GAT	TTT	TCT	CAC	CAG	CGA	AGG	CAT	TAC	AAA	GAA	TTT	CGA	1241
Arg	Asp	Lys	Asp	Phe	Ser	His	Gln	Arg	Arg	His	Tyr	Lys	Glu	Phe	Arg	
-150						-145					-140					
TTT	GAT	CTT	ACC	CAA	ATT	CCT	CAT	GGA	GAG	GCA	GTG	ACA	GCA	GCT	GAA	1289
Phe	Asp	Leu	Thr	Gln	Ile	Pro	His	Gly	Glu	Ala	Val	Thr	Ala	Ala	Glu	
-135				-130				-125						-120		
TTC	CGG	ATA	TAC	AAG	GAC	CGG	AGC	AAC	AAC	CGA	TTT	GAA	AAT	GAA	ACA	1337
Phe	Arg	Ile	Tyr	Lys	Asp	Arg	Ser	Asn	Asn	Arg	Phe	Glu	Asn	Glu	Thr	
				-115				-110						-105		
ATT	AAG	ATT	AGC	ATA	TAT	CAA	ATC	ATC	AAG	GAA	TAC	ACA	AAT	AGG	GAT	1385
Ile	Lys	Ile	Ser	Ile	Tyr	Gln	Ile	Ile	Lys	Glu	Tyr	Thr	Asn	Arg	Asp	
			-100				-95						-90			
GCA	GAT	CTG	TTC	TTG	TTA	GAC	ACA	AGA	AAG	GCC	CAA	GCT	TTA	GAT	GTG	1433
Ala	Asp	Leu	Phe	Leu	Leu	Asp	Thr	Arg	Lys	Ala	Gln	Ala	Leu	Asp	Val	
	-85					-80					-75					
GGT	TGG	CTT	GTC	TTT	GAT	ATC	ACT	GTG	ACC	AGC	AAT	CAT	TGG	GTG	ATT	1481
Gly	Trp	Leu	Val	Phe	Asp	Ile	Thr	Val	Thr	Ser	Asn	His	Trp	Val	Ile	
-70					-65						-60					
AAT	CCC	CAG	AAT	AAT	TTG	GGC	TTA	CAG	CTC	TGT	GCA	GAA	ACA	GGG	GAT	1529
Asn	Pro	Gln	Asn	Asn	Leu	Gly	Leu	Gln	Leu	Cys	Ala	Glu	Thr	Gly	Asp	
-55				-50				-45						-40		
GGA	CGC	AGT	ATC	AAC	GTA	AAA	TCT	GCT	GGT	CTT	GTG	GGA	AGA	CAG	GGA	1577
Gly	Arg	Ser	Ile	Asn	Val	Lys	Ser	Ala	Gly	Leu	Val	Gly	Arg	Gln	Gly	
			-35				-30							-25		
CCT	CAG	TCA	AAA	CAA	CCA	TTC	ATG	GTG	GCC	TTC	TTC	AAG	GCG	AGT	GAG	1625
Pro	Gln	Ser	Lys	Gln	Pro	Phe	Met	Val	Ala	Phe	Phe	Lys	Ala	Ser	Glu	
			-20				-15						-10			
GTA	CTT	CTT	CGA	TCC	GTG	AGA	GCA	GCC	AAC	AAA	CGA	AAA	AAT	CAA	AAC	1673
Val	Leu	Leu	Arg	Ser	Val	Arg	Ala	Ala	Asn	Lys	Arg	Lys	Asn	Gln	Asn	
	-5						1				5					
CGC	AAT	AAA	TCC	AGC	TCT	CAT	CAG	GAC	TCC	TCC	AGA	ATG	TCC	AGT	GTT	1721

Arg 10	Asn	Lys	Ser	Ser	Ser 15	His	Gln	Asp	Ser	Ser 20	Arg	Met	Ser	Ser	Val 25	
GGA Gly	GAT Asp	TAT Tyr	AAC Asn	ACA Thr	AGT Ser	GAG Glu	CAA Gln	AAA Lys	CAA Gln	GCC Ala	TGT Cys	AAG Lys	AAG Lys	CAC His	GAA Glu	1769
CTC Leu	TAT Tyr	GTG Val	AGC Ser	TTC Phe	CGG Arg	GAT Asp	CTG Leu	GGA Gly	TGG Trp	CAG Gln	GAC Asp	TGG Trp	ATT Ile	ATA Ile	GCA Ala	1817
CCA Pro	GAA Glu	GGA Gly	TAC Tyr	GCT Ala	GCA Ala	TTT Phe	TAT Tyr	TGT Cys	GAT Asp	GGA Gly	GAA Glu	TGT Cys	TCT Ser	TTT Phe	CCA Pro	1865
CTT Leu	AAC Asn	GCC Ala	CAT His	ATG Met	AAT Asn	GCC Ala	ACC Thr	AAC Asn	CAC His	GCT Ala	ATA Ile	GTT Val	CAG Gln	ACT Thr	CTG Leu	1913
GTT Val	CAT His	CTG Leu	ATG Met	TTT Phe	CCT Pro	GAC Asp	CAC His	GTA Val	CCA Pro	AAG Lys	CCT Pro	TGT Cys	TGT Cys	GCT Ala	CCA Pro	1961
ACC Thr	AAA Lys	TTA Leu	AAT Asn	GCC Ala	ATC Ile	TCT Ser	GTT Val	CTG Leu	TAC Tyr	TTT Phe	GAT Asp	GAC Asp	AGC Ser	TCC Ser	AAT Asn	2009
GTC Val	ATT Ile	TTG Leu	AAA Lys	AAA Lys	TAT Tyr	AGA Arg	AAT Asn	ATG Met	GTA Val	GTA Val	CGC Arg	TCA Ser	TGT Cys	GGC Gly	TGC Cys	2057
CAC His	TAATATTAAA	TAATATTGAT	AATAACAAAA	AGATCTGTAT	TAAGGTTTAT											2110
GGCTGCAATA	AAAAGCATAC	TTTCAGACAA	ACAGAAAAAA	AAA												2153

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 454 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met -316	His -315	Leu	Thr	Val	Phe	Leu	Leu	Lys	Gly	Ile	Val	Gly	Phe	Leu	Trp	
Ser -300	Cys	Trp	Val	Leu	Val	Gly	Tyr	Ala	Lys	Gly	Gly	Leu	Gly	Asp	Asn	-285
His	Val	His	Ser	Ser	Phe	Ile	Tyr	Arg	Arg	Leu	Arg	Asn	His	Glu	Arg	
Arg	Glu	Ile	Gln	Arg	Glu	Ile	Leu	Ser	Ile	Leu	Gly	Leu	Pro	His	Arg	

Pro Arg Pro Phe Ser Pro Gly Lys Met Thr Asn Gln Ala Ser Ser Ala
 -250 -245 -240

Pro Leu Phe Met Leu Asp Leu Tyr Asn Ala Glu Glu Asn Pro Glu Glu
 -235 -230 -225

Ser Glu Tyr Ser Val Arg Ala Ser Leu Ala Glu Glu Thr Arg Gly Ala
 -220 -215 -210 -205

Arg Lys Gly Tyr Pro Ala Ser Pro Asn Gly Tyr Pro Arg Arg Ile Gln
 -200 -195 -190

Leu Ser Arg Thr Thr Pro Leu Thr Thr Gln Ser Pro Pro Leu Ala Ser
 -185 -180 -175

Leu His Asp Thr Asn Phe Leu Asn Asp Ala Asp Met Val Met Ser Phe
 -170 -165 -160

Val Asn Leu Val Glu Arg Asp Lys Asp Phe Ser His Gln Arg Arg His
 -155 -150 -145

Tyr Lys Glu Phe Arg Phe Asp Leu Thr Gln Ile Pro His Gly Glu Ala
 -140 -135 -130 -125

Val Thr Ala Ala Glu Phe Arg Ile Tyr Lys Asp Arg Ser Asn Asn Arg
 -120 -115 -110

Phe Glu Asn Glu Thr Ile Lys Ile Ser Ile Tyr Gln Ile Ile Lys Glu
 -105 -100 -95

Tyr Thr Asn Arg Asp Ala Asp Leu Phe Leu Leu Asp Thr Arg Lys Ala
 -90 -85 -80

Gln Ala Leu Asp Val Gly Trp Leu Val Phe Asp Ile Thr Val Thr Ser
 -75 -70 -65

Asn His Trp Val Ile Asn Pro Gln Asn Asn Leu Gly Leu Gln Leu Cys
 -60 -55 -50 -45

Ala Glu Thr Gly Asp Gly Arg Ser Ile Asn Val Lys Ser Ala Gly Leu
 -40 -35 -30

Val Gly Arg Gln Gly Pro Gln Ser Lys Gln Pro Phe Met Val Ala Phe
 -25 -20 -15

Phe Lys Ala Ser Glu Val Leu Leu Arg Ser Val Arg Ala Ala Asn Lys
 -10 -5 1

Arg Lys Asn Gln Asn Arg Asn Lys Ser Ser Ser His Gln Asp Ser Ser
 5 10 15 20

Arg Met Ser Ser Val Gly Asp Tyr Asn Thr Ser Glu Gln Lys Gln Ala
 25 30 35

Cys Lys Lys His Glu Leu Tyr Val Ser Phe Arg Asp Leu Gly Trp Gln
 40 45 50

Asp Trp Ile Ile Ala Pro Glu Gly Tyr Ala Ala Phe Tyr Cys Asp Gly
 55 60 65

Glu Cys Ser Phe Pro Leu Asn Ala His Met Asn Ala Thr Asn His Ala
 70 75 80
 Ile Val Gln Thr Leu Val His Leu Met Phe Pro Asp His Val Pro Lys
 85 90 95 100
 Pro Cys Cys Ala Pro Thr Lys Leu Asn Ala Ile Ser Val Leu Tyr Phe
 105 110 115
 Asp Asp Ser Ser Asn Val Ile Leu Lys Lys Tyr Arg Asn Met Val Val
 120 125 130
 Arg Ser Cys Gly Cys His
 135

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1003 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo sapiens
 - (F) TISSUE TYPE: Human Heart
- (vii) IMMEDIATE SOURCE:
 - (A) LIBRARY: Human heart cDNA library stratagene catalog
#936208
 - (B) CLONE: hH38
- (viii) POSITION IN GENOME:
 - (C) UNITS: bp
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 8..850
- (ix) FEATURE:
 - (A) NAME/KEY: mat_peptide
 - (B) LOCATION: 427..843
- (ix) FEATURE:
 - (A) NAME/KEY: mRNA
 - (B) LOCATION: 1..997
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GAATTCC GAG CCC CAT TGG AAG GAG TTC CGC TTT GAC CTG ACC CAG ATC
 Glu Pro His Trp Lys Glu Phe Arg Phe Asp Leu Thr Gln Ile
 -139 -135 -130

CCG Pro -125	GCT Ala	GGG Gly	GAG Glu	GCG Ala	GTC Val -120	ACG Thr	GCT Ala	GCG Ala	GAG Glu -115	TTC Phe	CGG Arg	ATT Ile	TAC Tyr	AAG Lys	GTG Val -110	97
CCC Pro	AGC Ser	ATC Ile	CAC His	CTG Leu -105	CTC Leu	AAC Asn	AGG Arg	ACC Thr	CTC Leu -100	CAC His	GTC Val	AGC Ser	ATG Met	TTC Phe -95	CAG Gln	145
GTG Val	GTC Val	CAG Gln -90	GAG Glu -90	CAG Gln	TCC Ser	AAC Asn	AGG Arg	GAG Glu -85	TCT Ser	GAC Asp	TTG Leu	TTC Phe	TTT Phe -80	TTG Leu	GAT Asp	193
CTT Leu	CAG Gln -75	ACG Thr	CTC Leu	CGA Arg	GCT Ala	GGA Gly	GAC Asp -70	GAG Glu	GGC Gly	TGG Trp	CTG Leu	GTG Val -65	CTG Leu	GAT Asp	GTC Val	241
ACA Thr -60	GCA Ala	GCC Ala	AGT Ser	GAC Asp	TGC Cys	TGG Trp -55	TTG Leu	CTG Leu	AAG Lys	CGT Arg	CAC His -50	AAG Lys	GAC Asp	CTG Leu	GGA Gly	289
CTC Leu -45	CGC Arg	CTC Leu	TAT Tyr	GTG Val -40	GAG Glu -40	ACT Thr	GAG Glu	GAT Asp	GGG Gly	CAC His -35	AGC Ser	GTG Val	GAT Asp	CCT Pro	GGC Gly -30	337
CTG Leu	GCC Ala	GGC Gly	CTG Leu -25	CTG Leu -25	GGT Gly	CAA Gln	CGG Arg	GCC Ala -20	CCA Pro -20	CGC Arg	TCC Ser	CAA Gln	CAG Gln -15	CCT Pro -15	TTC Phe	385
GTG Val	GTC Val	ACT Thr -10	TTC Phe -10	TTC Phe	AGG Arg	GCC Ala	AGT Ser	CCG Pro -5	AGT Ser	CCC Pro	ATC Ile	CGC Arg	ACC Thr 1	CCT Pro	CGG Arg	433
GCA Ala 5	GTG Val	AGG Arg	CCA Pro	CTG Leu	AGG Arg	AGG Arg 10	AGG Arg	CAG Gln	CCG Pro	AAG Lys	AAA Lys 15	AGC Ser	AAC Asn	GAG Glu	CTG Leu	481
CCG Pro 20	CAG Gln	GCC Ala	AAC Asn	CGA Arg	CTC Leu 25	CCA Pro	GGG Gly	ATC Ile	TTT Phe	GAT Asp 30	GAC Asp	GTC Val	CAC His	GGC Gly	TCC Ser 35	529
CAC His	GGC Gly	CGG Arg	CAG Gln	GTC Val 40	TGC Cys	CGT Arg	CGG Arg	CAC His	GAG Glu 45	CTC Leu	TAC Tyr	GTC Val	AGC Ser	TTC Phe 50	CAG Gln	577
GAC Asp	CTT Leu	GGC Gly	TGG Trp 55	CTG Leu	GAC Asp	TGG Trp	GTC Val	ATC Ile 60	GCC Ala	CCC Pro	CAA Gln	GGC Gly	TAC Tyr 65	TCA Ser	GCC Ala	625
TAT Tyr	TAC Tyr	TGT Cys 70	GAG Glu	GGG Gly	GAG Glu	TGC Cys	TCC Ser 75	TTC Phe	CCG Pro	CTG Leu	GAC Asp	TCC Ser 80	TGC Cys	ATG Met	AAC Asn	673
GCC Ala 85	ACC Thr	AAC Asn	CAC His	GCC Ala	ATC Ile	CTG Leu 90	CAG Gln	TCC Ser	CTG Leu	GTG Val	CAC His 95	CTG Leu	ATG Met	AAG Lys	CCA Pro	721
AAC Asn 100	GCA Ala	GTC Val	CCC Pro	AAG Lys	GCG Ala 105	TGC Cys	TGT Cys	GCA Ala	CCC Pro	ACC Thr 110	AAG Lys	CTG Leu	AGC Ser	GCC Ala	ACC Thr 115	769

TCT GTG CTC TAC TAT GAC ACC AGC AAC AAC GTC ATC CTG CCG AAG CAC	817
Ser Val Leu Tyr Tyr Asp Ser Ser Asn Asn Val Ile Leu Arg Lys His	
120 125 130	
CGC AAC ATG GTG GTC AAG GCC TGC GGC TGC CAC TGAGTCAGCC CGCCCAGCCC	870
Arg Asn Met Val Val Lys Ala Cys Gly Cys His	
135 140	
TACTGCAGCC ACCCTTCTCA TCTGGATCGG GCCCTGCAGA GGCAGAAAAC CCTTAAATGC	930
TGTCACAGCT CAAGCAGGAG TGTCAGGGGC CCTCACTCTC GGTGCCTACT TCCTGTCAGG	990
CTTCTGGGAA TTC	1003

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 281 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Glu Pro His Trp Lys Glu Phe Arg Phe Asp Leu Thr Gln Ile Pro Ala	
-139 -135 -130 -125	
Gly Glu Ala Val Thr Ala Ala Glu Phe Arg Ile Tyr Lys Val Pro Ser	
-120 -115 -110	
Ile His Leu Leu Asn Arg Thr Leu His Val Ser Met Phe Gln Val Val	
-105 -100 -95	
Gln Glu Gln Ser Asn Arg Glu Ser Asp Leu Phe Phe Leu Asp Leu Gln	
-90 -85 -80	
Thr Leu Arg Ala Gly Asp Glu Gly Trp Leu Val Leu Asp Val Thr Ala	
-75 -70 -65 -60	
Ala Ser Asp Cys Trp Leu Leu Lys Arg His Lys Asp Leu Gly Leu Arg	
-55 -50 -45	
Leu Tyr Val Glu Thr Glu Asp Gly His Ser Val Asp Pro Gly Leu Ala	
-40 -35 -30	
Gly Leu Leu Gly Gln Arg Ala Pro Arg Ser Gln Gln Pro Phe Val Val	
-25 -20 -15	
Thr Phe Phe Arg Ala Ser Pro Ser Pro Ile Arg Thr Pro Arg Ala Val	
-10 -5 1 5	
Arg Pro Leu Arg Arg Arg Gln Pro Lys Lys Ser Asn Glu Leu Pro Gln	
10 15 20	
Ala Asn Arg Leu Pro Gly Ile Phe Asp Asp Val His Gly Ser His Gly	
25 30 35	
Arg Gln Val Cys Arg Arg His Glu Leu Tyr Val Ser Phe Gln Asp Leu	

Gly Trp Leu Asp Trp Val Ile Ala Pro Gln Gly Tyr Ser Ala Tyr Tyr
 55 60 65
 Cys Glu Gly Glu Cys Ser Phe Pro Leu Asp Ser Cys Met Asn Ala Thr
 70 75 80 85
 Asn His Ala Ile Leu Gln Ser Leu Val His Leu Met Lys Pro Asn Ala
 90 95 100
 Val Pro Lys Ala Cys Cys Ala Pro Thr Lys Leu Ser Ala Thr Ser Val
 105 110 115
 Leu Tyr Tyr Asp Ser Ser Asn Asn Val Ile Leu Arg Lys His Arg Asn
 120 125 130
 Met Val Val Lys Ala Cys Gly Cys His
 135 140

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3623 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vii) IMMEDIATE SOURCE:

- (B) CLONE: pALBP2-781

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2724..3071

(ix) FEATURE:

- (A) NAME/KEY: terminator
- (B) LOCATION: 3150..3218

(ix) FEATURE:

- (A) NAME/KEY: RBS
- (B) LOCATION: 2222..2723

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GACGAAAGGG CCTCGTGATA CGCCTATTTT TATAGGTAA TGTCATGATA ATAATGGTTT	60
CTTAGACGTC AGGTGGCACT TTTCGGGGAA ATGTGCGCGG AACCCCTATT TGTTTATTTT	120
TCTAAATACA TTCAAATATG TATCCGCTCA TGAGACAATA ACCCTGATAA ATGCTTCAAT	180
AATATTGAAA AAGGAAGAGT ATGAGTATTC AACATTTCCG TGTCGCCCTT ATTCCCTTTT	240
TTGCGGCATT TTGCCTTCCT GTTTTGTGCTC ACCCAGAAAC GCTGGTGAAA GTAAAAGATG	300
CTGAAGATCA GTTGGGTGCA CGAGTGGGTT ACATCGAACT GGATCTCAAC AGCGGTAAGA	360

TCCTTGAGAG	TTTTCGCCCC	GAAGAACGTT	TTCCAATGAT	GAGCACTTTT	AAAGTTCTGC	420
TATGTGGCGC	GGTATTATCC	CGTATTGACG	CCGGGCAAGA	GCAACTCGGT	CGCCGCATAC	480
ACTATTCTCA	GAATGACTTG	GTTGAGTACT	CACCAGTCAC	AGAAAAGCAT	CTTACGGATG	540
GCATGACAGT	AAGAGAATTA	TGCAGTGCTG	CCATAACCAT	GAGTGATAAC	ACTGCGGCCA	600
ACTTACTTCT	GACAACGATC	GGAGGACCGA	AGGAGCTAAC	CGCTTTTTTG	CACAACATGG	660
GGGATCATGT	AACTCGCCTT	GATCGTTGGG	AACCGGAGCT	GAATGAAGCC	ATACCAAACG	720
ACGAGCGTGA	CACCACGATG	CCTGTAGCAA	TGGCAACAAC	GTTGCGCAAA	CTATTAACTG	780
GCGAACTACT	TACTCTAGCT	TCCCGGCAAC	AATTAATAGA	CTGGATGGAG	GCGGATAAAG	840
TTGCAGGACC	ACTTCTGCGC	TCGGCCCTTC	CGGCTGGCTG	GTTTATTGCT	GATAAATCTG	900
GAGCCGGTGA	GCGTGGGTCT	CGCGGTATCA	TTGCAGCACT	GGGGCCAGAT	GGTAAGCCCT	960
CCCGTATCGT	AGTTATCTAC	ACGACGGGGA	GTCAGGCAAC	TATGGATGAA	CGAAATAGAC	1020
AGATCGCTGA	GATAGGTGCC	TCACTGATTA	AGCATTGGTA	ACTGTCAGAC	CAAGTTTACT	1080
CATATATACT	TTAGATTGAT	TTAAAACTTC	ATTTTTAATT	TAAAAGGATC	TAGGTGAAGA	1140
TCCTTTTTGA	TAATCTCATG	ACCAAAATCC	CTTAACGTGA	GTTTTCGTTC	CACTGAGCGT	1200
CAGACCCCGT	AGAAAAGATC	AAAGGATCTT	CTTGAGATCC	TTTTTTTTCTG	CGCGTAATCT	1260
GCTGCTTGCA	AACAAAAAAA	CCACCGCTAC	CAGCGGTGGT	TTGTTTGCCG	GATCAAGAGC	1320
TACCAACTCT	TTTTCCGAAG	GTAAGTGGCT	TCAGCAGAGC	GCAGATACCA	AATACTGTCC	1380
TTCTAGTGTA	GCCGTAGTTA	GGCCACCACT	TCAAGAACTC	TGTAGCACCG	CCTACATACC	1440
TCGCTCTGCT	AATCCTGTTA	CCAGTGGCTG	CTGCCAGTGG	CGATAAGTCG	TGTCTTACCG	1500
GGTTGGACTC	AAGACGATAG	TTACCGGATA	AGGCGCAGCG	GTCGGGCTGA	ACGGGGGGTT	1560
CGTGACACACA	GCCCAGCTTG	GAGCGAACGA	CCTACACCGA	ACTGAGATAC	CTACAGCGTG	1620
AGCATTGAGA	AAGCGCCACG	CTTCCCGAAG	GGAGAAAGGC	GGACAGGTAT	CCGGTAAGCG	1680
GCAGGGTCGG	AACAGGAGAG	CGCACGAGGG	AGCTTCCAGG	GGGAAACGCC	TGGTATCTTT	1740
ATAGTCCTGT	CGGGTTTCGC	CACCTCTGAC	TTGAGCGTCG	ATTTTTGTGA	TGCTCGTCAG	1800
GGGGGCGGAG	CCTATGGAAA	AACGCCAGCA	ACGCGGCCTT	TTACGGTTC	CTGGCCTTTT	1860
GCTGGCCTTT	TGCTCACATG	TTCTTTCCTG	CGTTATCCCC	TGATTCTGTG	GATAACCGTA	1920
TTACCGCCTT	TGAGTGAGCT	GATACCGCTC	GCCGCAGCCG	AACGACCGAG	CGCAGCGAGT	1980
CAGTGAGCGA	GGAAGCGGAA	GAGCGCCCAA	TACGCAAACC	GCCTCTCCCC	GCGCGTTGGC	2040
CGATTCATTA	ATGCAGAATT	GATCTCTCAC	CTACCAAACA	ATGCCCCCCT	GCAAAAAATA	2100
AATTCATATA	AAAAACATAC	AGATAACCAT	CTGCGGTGAT	AAATTATCTC	TGGCGGTGTT	2160

GACATAAATA	CCACTGGCGG	TGATACTGAG	CACATCAGCA	GGACGCACTG	ACCACCATGA	2220										
AGGTGACGCT	CTTAAAAATT	AAGCCCTGAA	GAAGGGCAGC	ATTCAAAGCA	GAAGGCTTTG	2280										
GGGTGTGTGA	TACGAAACGA	AGCATTGGCC	GTAAGTGCGA	TTCCGGATTA	GCTGCCAATG	2340										
TGCCAATCGC	GGGGGGTTTT	CGTTCAGGAC	TACAACTGCC	ACACACCACC	AAAGCTAACT	2400										
GACAGGAGAA	TCCAGATGGA	TGCACAAACA	CGCCGCCGCG	AACGTCGCGC	AGAGAAACAG	2460										
GCTCAATGGA	AAGCAGCAAA	TCCCCTGTTG	GTTGGGGTAA	GCGCAAAACC	AGTTCCGAAA	2520										
GATTTTTTTT	ACTATAAACG	CTGATGGAAG	CGTTTATGCG	GAAGAGGTAA	AGCCCTTCCC	2580										
GAGTAACAAA	AAAACAACAG	CATAAATAAC	CCCGCTCTTA	CACATTCCAG	CCCTGAAAAA	2640										
GGGCATCAAA	TTAAACCACA	CCTATGGTGT	ATGCATTTAT	TTGCATACAT	TCAATCAATT	2700										
GTTATCTAAG	GAAATACTTA	CAT	ATG	CAA	GCT	AAA	CAT	AAA	CAA	CGT	AAA	2750				
			Met	Gln	Ala	Lys	His	Lys	Gln	Arg	Lys					
			1					5								
CGT	CTG	AAA	TCT	AGC	TGT	AAG	AGA	CAC	CCT	TTG	TAC	GTG	GAC	TTC	AGT	2798
Arg	Leu	Lys	Ser	Ser	Cys	Lys	Arg	His	Pro	Leu	Tyr	Val	Asp	Phe	Ser	
10					15					20					25	
GAC	GTG	GGG	TGG	AAT	GAC	TGG	ATT	GTG	GCT	CCC	CCG	GGG	TAT	CAC	GCC	2846
Asp	Val	Gly	Trp	Asn	Asp	Trp	Ile	Val	Ala	Pro	Pro	Gly	Tyr	His	Ala	
				30					35					40		
TTT	TAC	TGC	CAC	GGA	GAA	TGC	CCT	TTT	CCT	CTG	GCT	GAT	CAT	CTG	AAC	2894
Phe	Tyr	Cys	His	Gly	Glu	Cys	Pro	Phe	Pro	Leu	Ala	Asp	His	Leu	Asn	
			45					50					55			
TCC	ACT	AAT	CAT	GCC	ATT	GTT	CAG	ACG	TTG	GTC	AAC	TCT	GTT	AAC	TCT	2942
Ser	Thr	Asn	His	Ala	Ile	Val	Gln	Thr	Leu	Val	Asn	Ser	Val	Asn	Ser	
		60					65					70				
AAG	ATT	CCT	AAG	GCA	TGC	TGT	GTC	CCG	ACA	GAA	CTC	AGT	GCT	ATC	TCG	2990
Lys	Ile	Pro	Lys	Ala	Cys	Cys	Val	Pro	Thr	Glu	Leu	Ser	Ala	Ile	Ser	
	75					80					85					
ATG	CTG	TAC	CTT	GAC	GAG	AAT	GAA	AAG	GTT	GTA	TTA	AAG	AAC	TAT	CAG	3038
Met	Leu	Tyr	Leu	Asp	Glu	Asn	Glu	Lys	Val	Val	Leu	Lys	Asn	Tyr	Gln	
90					95					100					105	
GAC	ATG	GTT	GTG	GAG	GGT	TGT	GGG	TGT	CGC	TAGTACAGCA	AAATTAAATA					3088
Asp	Met	Val	Val	Glu	Gly	Cys	Gly	Cys	Arg							
				110					115							
CATAAATATA	TATATATATA	TATATTTTAG	AAAAAAGAAA	AAAATCTAGA	GTCGACCTGC	3148										
AGTAATCGTA	CAGGGTAGTA	CAAATAAAAA	AGGCACGTCA	GATGACGTGC	CTTTTTTCTT	3208										
GTGAGCAGTA	AGCTTGGCAC	TGGCCGTCGT	TTTACAACGT	CGTGACTGGG	AAAACCCTGG	3268										
CGTTACCCAA	CTTAATCGCC	TTGCAGCACA	TCCCCCTTTC	GCCAGCTGGC	GTAATAGCGA	3328										
AGAGGCCCGC	ACCGATCGCC	CTTCCCAACA	GTTGCGCAGC	CTGAATGGCG	AATGGCGCCT	3388										

GATGCGGTAT TTTCTCCTTA CGCATCTGTG CGGTATTTCA CACCGCATAT ATGGTGCACT	3448
CTCAGTACAA TCTGCTCTGA TGCCGCATAG TTAAGCCAGC CCCGACACCC GCCAACACCC	3508
GCTGACGCGC CCTGACGGGC TTGTCTGCTC CCGGCATCCG CTTACAGACA AGCTGTGACC	3568
GTCTCCGGGA GCTGCATGTG TCAGAGGTTT TCACCGTCAT CACCGAAACG CGCGA	3623

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 115 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met	Gln	Ala	Lys	His	Lys	Gln	Arg	Lys	Arg	Leu	Lys	Ser	Ser	Cys	Lys	
1				5					10					15		
Arg	His	Pro	Leu	Tyr	Val	Asp	Phe	Ser	Asp	Val	Gly	Trp	Asn	Asp	Trp	
			20					25					30			
Ile	Val	Ala	Pro	Pro	Gly	Tyr	His	Ala	Phe	Tyr	Cys	His	Gly	Glu	Cys	
		35					40					45				
Pro	Phe	Pro	Leu	Ala	Asp	His	Leu	Asn	Ser	Thr	Asn	His	Ala	Ile	Val	
	50					55					60					
Gln	Thr	Leu	Val	Asn	Ser	Val	Asn	Ser	Lys	Ile	Pro	Lys	Ala	Cys	Cys	
65				70						75					80	
Val	Pro	Thr	Glu	Leu	Ser	Ala	Ile	Ser	Met	Leu	Tyr	Leu	Asp	Glu	Asn	
				85					90					95		
Glu	Lys	Val	Val	Leu	Lys	Asn	Tyr	Gln	Asp	Met	Val	Val	Glu	Gly	Cys	
			100					105					110			
Gly	Cys	Arg														
																115

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 14 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 41 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GAGGGTTGTG GGTGTCGCTA GTGAGTCGAC TACAGCAAAT T

41

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 38 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GGATGTGGGT GCCGCTGACT CTAGAGTCGA CGGAATTC

38

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 31 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

AATTCACCAT GATTCCTGGT AACCGAATGC T

31

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 25 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GTGGTACTAA GGACCATTTGG CTTAC

25

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CGACCTGCAG CCATGCATCT GACTGTA

27

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

TGCCTGCAGT TTAATATTAG TGGCAGC

27

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CGACCTGCAG CCACC

15

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 81 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

TCGACCCACC ATGCCGGGGC TGGGGCGGAG GGCGCAGTGG CTGTGCTGGT GGTGGGGGCT 60
GTGCTGCAGC TGCTGCGGGC C 81

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 73 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CGCAGCAGCT GCACAGCAGC CCCCACCACC AGCACAGCCA CTGCGCCCTC CGCCCCAGCC 60
CCGGCATGGT GGG 73

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

TCGACTGGTT T 11

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

(2) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 18 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

TCGACAGGCT CGCCTGCA

18

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GTCCGAGCGG

10

(2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 29 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

CAGGTCGACC CACCATGCAC GTGCGCTCA

29

(2) INFORMATION FOR SEQ ID NO:30:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 27 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

27

TCTGTCGACC TCGGAGGAGC TAGTGGC

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1794 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: bmp-3

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 321..1136

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

AGATCTTGAA AACACCCGGG CCACACACGC CGCGACCTAC AGCTCTTTCT CAGCGTTGGA	60
GTGGAGACGG CGCCCGCAGC GCCCTGCGCG GGTGAGGTCC GCGCAGCTGC TGGGGAAGAG	120
CCCACCTGTC AGGCTGCGCT GGGTCAGCGC AGCAAGTGGG GCTGGCCGCT ATCTCGCTGC	180
ACCCGGCCGC GTCCCGGGCT CCGTGCGCCC TCGCCCCAGC TGGTTTGGAG TTCAACCCTC	240
GGCTCCGCCG CCGGCTCCTT GCGCCTTCGG AGTGTCCCGC AGCGACGCCG GGAGCCGACG	300
CGCCGCGCGG GTACCTAGCC ATG GCT GGG GCG AGC AGG CTG CTC TTT CTG	350
Met Ala Gly Ala Ser Arg Leu Leu Phe Leu	
1 5 10	
TGG CTG GGC TGC TTC TGC GTG AGC CTG GCG CAG GGA GAG AGA CCG AAG	398
Trp Leu Gly Cys Phe Cys Val Ser Leu Ala Gln Gly Glu Arg Pro Lys	
15 20 25	
CCA CCT TTC CCG GAG CTC CGC AAA GCT GTG CCA GGT GAC CGC ACG GCA	446
Pro Pro Phe Pro Glu Leu Arg Lys Ala Val Pro Gly Asp Arg Thr Ala	
30 35 40	
GGT GGT GGC CCG GAC TCC GAG CTG CAG CCG CAA GAC AAG GTC TCT GAA	494
Gly Gly Gly Pro Asp Ser Glu Leu Gln Pro Gln Asp Lys Val Ser Glu	
45 50 55	
CAC ATG CTG CGG CTC TAT GAC AGG TAC AGC ACG GTC CAG GCG GCC CGG	542
His Met Leu Arg Leu Tyr Asp Arg Tyr Ser Thr Val Gln Ala Ala Arg	
60 65 70	
ACA CCG GGC TCC CTG GAG GGA GGC TCG CAG CCC TGG CGC CCT CGG CTC	590
Thr Pro Gly Ser Leu Glu Gly Gly Ser Gln Pro Trp Arg Pro Arg Leu	
75 80 85 90	
CTG CGC GAA GGC AAC ACG GTT CGC AGC TTT CGG GCG GCA GCA GCA GAA	638

Leu	Arg	Glu	Gly	Asn	Thr	Val	Arg	Ser	Phe	Arg	Ala	Ala	Ala	Ala	Glu	
				95					100						105	
ACT	CTT	GAA	AGA	AAA	GGA	CTG	TAT	ATC	TTC	AAT	CTG	ACA	TCG	CTA	ACC	686
Thr	Leu	Glu	Arg	Lys	Gly	Leu	Tyr	Ile	Phe	Asn	Leu	Thr	Ser	Leu	Thr	
			110					115					120			
AAG	TCT	GAA	AAC	ATT	TTG	TCT	GCC	ACA	CTG	TAT	TTC	TGT	ATT	GGA	GAG	734
Lys	Ser	Glu	Asn	Ile	Leu	Ser	Ala	Thr	Leu	Tyr	Phe	Cys	Ile	Gly	Glu	
			125				130					135				
CTA	GGA	AAC	ATC	AGC	CTG	AGT	TGT	CCA	GTG	TCT	GGA	GGA	TGC	TCC	CAT	782
Leu	Gly	Asn	Ile	Ser	Leu	Ser	Cys	Pro	Val	Ser	Gly	Gly	Cys	Ser	His	
	140					145					150					
CAT	GCT	CAG	AGG	AAA	CAC	ATT	CAG	ATT	GAT	CTT	TCT	GCA	TGG	ACC	CTC	830
His	Ala	Gln	Arg	Lys	His	Ile	Gln	Ile	Asp	Leu	Ser	Ala	Trp	Thr	Leu	
155					160				165						170	
AAA	TTC	AGC	AGA	AAC	CAA	AGT	CAA	CTC	CTT	GGC	CAT	CTG	TCA	GTG	GAT	878
Lys	Phe	Ser	Arg	Asn	Gln	Ser	Gln	Leu	Leu	Gly	His	Leu	Ser	Val	Asp	
				175				180						185		
ATG	GCC	AAA	TCT	CAT	CGA	GAT	ATT	ATG	TCC	TGG	CTG	TCT	AAA	GAT	ATC	926
Met	Ala	Lys	Ser	His	Arg	Asp	Ile	Met	Ser	Trp	Leu	Ser	Lys	Asp	Ile	
			190					195					200			
ACT	CAA	TTC	TTG	AGG	AAG	GCC	AAA	GAA	AAT	GAA	GAG	TTC	CTC	ATA	GGA	974
Thr	Gln	Phe	Leu	Arg	Lys	Ala	Lys	Glu	Asn	Glu	Glu	Phe	Leu	Ile	Gly	
		205					210					215				
TTT	AAC	ATT	ACG	TCC	AAG	GGA	CGC	CAG	CTG	CCA	AAG	AGG	AGG	TTA	CCT	1022
Phe	Asn	Ile	Thr	Ser	Lys	Gly	Arg	Gln	Leu	Pro	Lys	Arg	Arg	Leu	Pro	
	220					225					230					
TTT	CCA	GAG	CCT	TAT	ATC	TTG	GTA	TAT	GCC	AAT	GAT	GCC	GCC	ATT	TCT	1070
Phe	Pro	Glu	Pro	Tyr	Ile	Leu	Val	Tyr	Ala	Asn	Asp	Ala	Ala	Ile	Ser	
235					240				245						250	
GAG	CCA	GAA	AGT	GTG	GTA	TCA	AGC	TTA	CAG	GGA	CAC	CGG	AAT	TTT	CCC	1118
Glu	Pro	Glu	Ser	Val	Val	Ser	Ser	Leu	Gln	Gly	His	Arg	Asn	Phe	Pro	
				255					260					265		
ACT	GGA	ACT	GTT	CCC	AAA	TGGG	ATAGCC	ACATCAGAGC	TGCCCTTTCC							1166
Thr	Gly	Thr	Val	Pro	Lys											
			270													
ATTGAGCGGA	GGAAGAAGCG	CTCTACTGGG	GTCTTGCTGC	CTCTGCAGAA	CAACGAGCTT											1226
CCTGGGGCAG	AATACCAGTA	TAAAAAGGAT	GAGGTGTGGG	AGGAGAGAAA	GCCTTACAAG											1286
ACCCTTCAGG	CTCAGGCCCC	TGAAAAGAGT	AAGAATAAAA	AGAAACAGAG	AAAGGGGCCT											1346
CATCGGAAGA	GCCAGACGCT	CCAATTTGAT	GAGCAGACCC	TGAAAAAGGC	AAGGAGAAAG											1406
CAGTGGATTG	AACCTCGGAA	TTGCGCCAGG														

GTGAGAGCTG TGGGGGTCGT TCCTGGGATT CCTGAGCCTT GCTGTGTACC AGAAAAGATG	1646
TCCTCACTCA GTATTTTATT CTTTGATGAA AATAAGAATG TAGTGCTTAA AGTATACCCT	1706
AACATGACAG TAGAGTCTTG CGCTTGCAGA TAACCTGGCA AAGAACTCAT TTGAATGCTT	1766
AATTCAATCT CTAGAGTCGA CGGAATTC	1794

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 272 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Met	Ala	Gly	Ala	Ser	Arg	Leu	Leu	Phe	Leu	Trp	Leu	Gly	Cys	Phe	Cys	
1				5					10					15		
Val	Ser	Leu	Ala	Gln	Gly	Glu	Arg	Pro	Lys	Pro	Pro	Phe	Pro	Glu	Leu	
			20					25					30			
Arg	Lys	Ala	Val	Pro	Gly	Asp	Arg	Thr	Ala	Gly	Gly	Gly	Pro	Asp	Ser	
		35					40					45				
Glu	Leu	Gln	Pro	Gln	Asp	Lys	Val	Ser	Glu	His	Met	Leu	Arg	Leu	Tyr	
	50					55					60					
Asp	Arg	Tyr	Ser	Thr	Val	Gln	Ala	Ala	Arg	Thr	Pro	Gly	Ser	Leu	Glu	
65					70					75					80	
Gly	Gly	Ser	Gln	Pro	Trp	Arg	Pro	Arg	Leu	Leu	Arg	Glu	Gly	Asn	Thr	
				85					90					95		
Val	Arg	Ser	Phe	Arg	Ala	Ala	Ala	Ala	Glu	Thr	Leu	Glu	Arg	Lys	Gly	
			100					105					110			
Leu	Tyr	Ile	Phe	Asn	Leu	Thr	Ser	Leu	Thr	Lys	Ser	Glu	Asn	Ile	Leu	
		115					120					125				
Ser	Ala	Thr	Leu	Tyr	Phe	Cys	Ile	Gly	Glu	Leu	Gly	Asn	Ile	Ser	Leu	
						135					140					
Ser	Cys	Pro	Val	Ser	Gly	Gly	Cys	Ser	His	His	Ala	Gln	Arg	Lys	His	
145					150				155					160		
Ile	Gln	Ile	Asp	Leu	Ser	Ala	Trp	Thr	Leu	Lys	Phe	Ser	Arg	Asn	Gln	
			165						170					175		
Ser	Gln	Leu	Leu	Gly	His	Leu	Ser	Val	Asp	Met	Ala	Lys	Ser	His	Arg	
			180					185					190			
Asp	Ile	Met	Ser	Trp	Leu	Ser	Lys	Asp	Ile	Thr	Gln	Phe	Leu	Arg	Lys	
		195					200					205				

Ala Lys Glu Asn Glu Glu Phe Leu Ile Gly Phe Asn Ile Thr Ser Lys
 210 215 220

Gly Arg Gln Leu Pro Lys Arg Arg Leu Pro Phe Pro Glu Pro Tyr Ile
 225 230 235 240

Leu Val Tyr Ala Asn Asp Ala Ala Ile Ser Glu Pro Glu Ser Val Val
 245 250 255

Ser Ser Leu Gln Gly His Arg Asn Phe Pro Thr Gly Thr Val Pro Lys
 260 265 270

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 449 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

ATGTCTCATA ATCGTTCTAA AACTCCAAAA AATCAAGAAG CTCTGCGTAT GGCCAACGTG	60
GCAGAGAACA GCAGCAGCGA CCAGAGGCAG GCCTGTAAGA AGCACGAGCT GTATGTCAGC	120
TTCCGAGACC TGGGCTGGCA GGA CTGGATC ATCGCGCCTG AAGGCTACGC CGCCTACTAC	180
TGTGAGGGGG AGTGTGCCTT CCCTCTGAAC TCCTACATGA ACGCCACCAA CCACGCCATC	240
GTGCAGACGC TGGTCCACTT CATCAACCCG GAAACGGTGC CCAAGCCCTG CTGTGCGCCC	300
ACGCAGCTCA ATGCCATCTC CGTCCTCTAC TTCGATGACA GCTCCAACGT CATCCTGAAG	360
AAATACAGAA ACATGGTGGT CCGGGCCTGT GGCTGCCACT AGCTCCTCCG AGAATTCAGA	420
CCCTTTGGGG CCAAGTTTTT CTGGATCCT	449

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

CAAGAAGGAG ATATACAT

18

(2) INFORMATION FOR SEQ ID NO:35:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 377 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

ATGCGTAAAC AATGGATTGA ACCACGTAAC TGTGCTCGTC GTTATCTGAA AGTAGACTTT	60
GCAGATATTG GCTGGAGTGA ATGGATTATC TCCCCCAAGT CCTTTGATGC CTATTATTGC	120
TCTGGAGCAT GCCAGTTCCC CATGCCAAAG TCTTTGAAGC CATCAAATCA TGCTACCATC	180
CAGAGTATAG TGAGAGCTGT GGGGGTCGTT CCTGGGATTC CTGAGCCTTG CTGTGTACCA	240
GAAAAGATGT CCTCACTCAG TATTTTATTC TTTGATGAAA ATAAGAATGT AGTGCTTAAA	300
GTATACCCTA ACATGACAGT AGAGTCTTGC GCTTGCAGAT AACCTGGCAA AGAACTCATT	360
TGAATGCTTA ATTCAAT	377